STATISTICS



كامل فليفل

جامعة البلقاء

كلية الهندسة التكنولوجية

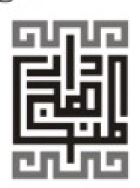
فتحي حمدان

جامعة البترا

كلية العلوم الإدارية والمالية



دار المناهج للنشر والتوزيع



دار المناهج للنشر والتوزيع
عمان /الأردن / شارع الملك حسين
بنايــة الشركـــة المتحدة للتأمين
هاتف 4650624 فاكس 4650664
ص.ب / 215308عمان 11122 الأردن

Dar Al-Manahej
Publishers & Distributor
Tel : (00962 6) 4650624
fax: 009626 4650664
Amman-King Hussein St.
P.O.Box: 215308 Amman 11122 Jordan
www.daralmanahej.com
info@daralmanahej.com
manahej9@hotmail.com
fayiz@daralmanahej.com

المملكة الأردنية الهاشمية
رقم الإيداع لدى دائـرة المكتبات والوثائق الوطنية
2009/2/430

519.5
فليفل، كامل
الإحصاء/ كامل فليفل، فتحي حمدان
عمان: دار المناهج ، 2009
( ) ص
ر.إ: 430/ 2/ 2009
الواصفات:الإحصاء الوصفي
* تم إعداد بيانات الفهرسة والتصنيف الأولية من قبل دائرة المكتبة الوطنية

المحتويات

الوحدة الثالثة

مقاييس التشتت

Measures of Dispersion

الوحدة الرابعة

الارتباط والانحدار

Correlation and Regression

الوحدة الخامسة

نظرية الاحتمالات

Probability Theory

الوحدة السادسة

التوزيعات الاحتمالية

Probability Distributions

الوحدة السابعة

التقدير واختبار الفرضيات

Estimation and Testing Hypothesis

الوحدة الثامنة

الأرقام القياسية

Index Numbers

الوحدة التاسعة

السلاسل الزمنية

Time Services

مُقَدِّمَة

نحمد الله على نعمه وفضله كما ينبغي لجلال وجهه وعظيم سلطانه والصلاة والسلام على الصادق الأمين محمد بن عبدالله وبعد:

فهذه الطبعة الثالثة من كتابنا الإحصاء نضعها بين يدي طلابنا الأعزاء راعينا فيها تطوير المعلومات وتحديثها بشكل ميسر في كل وحداته التي تغطي حاجة الطلاب، حيث تم كتابة المعادلات والرموز باللغة الإنجليزية. وقد تكوَّن الكتاب من تسع وحدات هي:

الوحدة الأولى : تعالج موضوع عرض البيانات الإحصائية وتبويبها.

الوحدة الثانية: تعالج مقاييس النزعة المركزية،

الوحدةالثالثة : تختص بمقاييس التشتت.

الوحدة الرابعة تعرض موضوع الارتباط والانحدار،

الوحدة الخامسة: تعالج موضوع الاحتمالات

الوحدة السادسة: تتعلق بموضوع التوزيعات الاحتمالية.

الوحدة السابعة : التقدير واختبار الفرضيات.

الوحدة الثامنة : تهتم بموضوع الأرقام القياسية.

الوحدة التاسعة : تهتم بالسلاسل الزمنية

وقد حرصنا في هذه الطبعة على وضع مسائل وأمثلة متنوعة تناسب كافة مستويات الطلبة في الجامعات.

ونعود ونؤكد على إخواننا المدرسين وكذلك أعزائنا الطلبة التكرم علينا بإبداء ملاحظاتهم وأفكارهم عن فقرات هذا الكتاب للاستفادة منها في الطبعات القادمة إن شاء الله ولهم منا جزيل الشكر والعرفان.

وفي الختام لا يسعنا إلا أن نشكر كل من ساهم في إخراج هذا الكتاب للقارئ الكريم.

والله ولي التوفيق

المؤلفان



الوحدة الأولى

جمع البيانات وتبويبها

Collecting and Organizing Data

جمع البيانات وتبويبها

Collecting and Organizing Data

مقدِّمة

بُنِي علم الإحصاء على مجموعة عناصر أساسية نجملها في هذا التعريف.

تعريف: علم الإحصاء هو مجموعة النظريات والطرق العلمية التي تبحث في جمع البيانات وعرضها وتحليلها واستخدام النتائج في التنبؤ أو التقرير واتخاذ القرارات بناء على ذلك.

يقسم علم الإحصاء إلى قسمين:

1- الإحصاء الوصفي: الذي يهتم بجمع البيانات الإحصائية وتبويبها فقط.

2- الإحصاء الاستدلالي: ويهتم في اتخاذ القرارات المبنية على النتائج المستخرجة من البيانات التي جمعت.

الطريقة الإحصائية Statistical Investigation

تعريف: الطريقة الإحصائية بأنها الخطوات المتبعة في عمل أي دراسة أو بحث إحصائي.

وهذه الخطوات هي:

أ- جمع البيانات الإحصائية:

وهي قيم المشاهدات للتجارب التي يجريها الباحث وكلما كانت دقيقة كلما كانت القرارات المتخذة بصددها أكثر صدقاً، وهناك طريقتين لجمع البيانات الإحصائية هي:

1- طريقة المسح الشامل.

2- طريقة العينة.

ب- تنظيم وعرض البيانات:

بعد جمع البيانات يقوم الباحث في وضع هذه البيانات في جداول مناسبة أو عرضها في رسوم بيانية أو أشكال هندسية أو توزيعات تكرارية.

جـ- تحليل البيانات:

وهي معالجة البيانات باستعمال العلاقات الرياضية واستخراج قيم واقترانات معينة تعبر عن هذه البيانات.

د- استقراء النتائج واتخاذ القرارات:

وهي إصدار الأحكام أو عمل الاستنتاجات الإحصائية حول المجتمع الإحصائي في ضوء النتائج المستخرجة.

أساليب جمع البيانات:

تجمع البيانات بأساليب عدة منها:

1- الأسلوب المباشر: عن طريق الميدان مباشرة.
2- الأسلوب غير المباشر: عن طريق السجلات أو الوثائق التاريخية.
3- أسلوب الاستبيان: وهي رزمة من الأوراق تحتوي على مجموعة من الأسئلة والاستفسارات تعبئ من قبل الشخص الخاضع للبحث.
4- أسلوب المقابلات الشخصية: وهي السؤال المباشر من قبل الباحث.
5- أسلوب الاختبارات الخاصة: أسلوب خاص يستخدم في حالات محددة فقط مثل اختبارات الذكاء.

العينة وطرق اختيارها Sampling

تعريف: العينة مجموعة جزئية من المجتمع الإحصائي.

تؤخذ العينة بعدة طرق حتى تكون ممثلة للمجتمع تمثيلاً صادقاً وهذا يتطلب أمرين:

1- تحديد هدف الدراسة: ويحدد الهدف بطرح أسئلة مثل لماذا نأخذ العينة؟ ما الذي نريده منها...الخ.
2- تحديد المجتمع الإحصائي: والمجتمع الإحصائي هو مجموعة كل العناصر قيد الدراسة ويسمى أحياناً مجتمع الهدف. وتسمى المجموعة التي تؤخذ منها العينة بمجتمع العينة. ونلاحظ أن مجتمع العينة جزء من مجتمع الهدف.

مثال:

إذا أردنا دراسة الصعوبات التي تواجه طلبة البرنامج التجاري في كليات المجتمع في مادة الإحصاء، فيكون مجتمع الهدف: هو جميع طلبة البرنامج التجاري في كليات المجتمع.

ومجتمع العينة يكون الكليات التي تؤخذ منها العينة مباشرة.

تعريف: حجم المجتمع (العينة) هو عدد عناصر المجتمع (العينة).

طرق اختيار العينة

1- طريقة العينة العشوائية البسيطة Simple Random sample

وهي أي عينة بحجم معين لها نفس الاحتمال. ويتم اختيارها بالطريقة التالية:

أ- إذا كانت العينة صغيرة (أقل من 30 مشاهدة) نعطي المشاهدات بطاقات متشابهة مرقمة ترقيماً متسلسلاً من (1) لغاية (n) حيث n حجم المجتمع، ثم نسحب بطاقات عشوائية بحجم العينة التي نريد.

ب- إذا كانت العينة كبيرة (أكبر من 30 مشاهدة) نعطي المشاهدات أرقاماً متسلسلة من (صفر) لغاية (n-1) "حيث n حجم المجتمع" بنفس العدد من الخانات (عدد خانات (n-1)). ثم نختار من جدول الأرقام العشوائية المرفق في نهاية الكتاب أرقاماً بحجم العينة شريطة أن تكون الأرقام أقل من (n-1). وسنوضح ذلك بالمثال اللاحق.

ملاحظة: يمكن استخدام هذه الطريقة إذا كان المجتمع صغيراً.

مثال:

إذا أردنا اختيار عينة مكونة (10) طلاب من مجتمع مكون من (900) طالب، نتبع الخطوات التالية:

أ- نعطي الطلاب أرقاماً متسلسلة من (000) ولغاية (899).

ب- نختار عشرة أرقـام مـن جـدول الأرقـام العـشوائية ونبـدأ مـن اليـسار ونتجـه عموديا للأسفل فإذا كـان الـرقم أقـل أو يـساوي (899) نقبلـه وبغـير ذلـك نرفضه، فتكون العينة مكونة من الأرقام التالية:

517, 540, 459, 35, 649, 156, 216, 505, 71, 279

ملاحظة: عند اختيار أرقاماً عشوائية من جدول الأرقام العشوائية يمكننا البـدء مـن أي مكان من الجـدول سـواء أفقيـاً أو عموديـاً ولكننـا اتبعنـا الأسـلوب الـسابق لتوحيـد الإجابات بين الطلبة.

2- طريقة العينة الطبقية Stratified sample

يقسم المجتمع الإحصائي إلى طبقات حسب صفات معينة ثم نختار عينـة عـشوائية بسيطة من كل طبقة من هذه الطبقات بنسبة حجم كل طبقة.

وتعطى النسبة بالقاعدة التالية:

$$\text{عدد العناصر الممثلة للطبقة في العينة} = \frac{\text{حجم الطبقة}}{\text{حجم المجتمع}} \times \text{حجم العينة}$$

مثال:

اختار عينة مكونـة مـن (20) طالـب مـن مجتمـع الجامعـة الأردنيـة المكـون مـن (1000) طالب منهم (400) طالب سنة أولى (300) طالب سنة ثانيـة (200) طالـب سنة ثالثة و (100) طالب سنة رابعة.

الحل:

$$\text{عدد عناصر طلاب السنة الأولى في العينة} = \frac{400}{1000} \times 20 = 8 \text{ طلاب}$$

$$\text{عدد عناصر طلاب السنة الثانية في العينة} = \frac{300}{1000} \times 20 = 6 \text{ طلاب}$$

$$\text{عدد عناصر طلاب السنة الثالثة في العينة} = \frac{200}{1000} \times 20 = 4 \text{ طلاب}$$

$$\text{عدد عناصر طلاب السنة الرابعة في العينة} = \frac{100}{1000} \times 20 = \text{ طالبان}$$

3- طريقة العينة العنقودية Cluster Sample

يقسم المجتمع الإحصائي إلى مجموعات جزئية واضحة تسمى كل منها طبقة. ثم نقسم الطبقة إلى طبقات أخرى وهكذا، ونختار عينة عشوائية بسيطة من الطبقة الأخيرة تتناسب مع حجم الطبقة.

مثال:

إذا أردنا دراسة فرص العمل لطلاب الجامعة الأردنية بعد التخرج.

نقوم في البداية بتقسيم الجامعة إلى كليات مثل كلية الطب، الهندسة، العلوم، التجارة،...الخ، ثم نقوم بتقسيم هذه الكليات إلى تخصصات ونأخذ عينة عشوائية بسيطة من كل تخصص ونجري الدراسة عليها.

4- طريقة العينة العشوائية المنتظمة Systematic Random sample

نختار الأرقام بهذه الطريقة بصورة منتظمة. بحيث يكون الفرق بين أي اختيارين متتالين يساوي مقداراً ثابتاً.

مثال:

إذا أردنا أن نجري دراسة على شارع مكون من (50) عمارة، وأردنا اختيار عينة مكونة من (5) عمارات فإننا نختار كل عاشر عمارة.

فمثلاً لو اخترنا العمارة رقم 7 تكون الثانية رقم 17 والثالثة رقم 27، والرابعة رقم 37، والخامسة رقم 47.

5- طريقة العينة المعيارية Standard sample

وهي العينة التي تتفق مع المجتمع الإحصائي في المقاييس الإحصائية فيكون لهما نفس الوسط والوسيط والانحراف المعياري. وتكون أكثر صدقا في تمثيل المجتمع من الطرق الأخرى.

مثال:

إذا أراد مصنع للأدوية دراسة مدى فعالية دواء ما لشفاء مرض معـين فإنـه يطبـق هذا الدواء على أول (10) مرضى ثم يرى مدى فعاليته ثم أول (20) مريض ثم يـرى مدى فعاليته ثم أول (30) مريض..الخ إلى أن تثبت فعالية الدواء فـيعمم هـذا الـدواء لعلاج المرض.

طرق عرض البيانات الإحصائية Graphic Presentation of data

يتم عرض البيانات بطرق سهلة وواضحة ليسهل على الباحث دراسة ظاهرة مـا. ويتم العرض بالطرق التالية:

1- طريقة الجدول Table method

وهي عرض الظاهرة مع مسمى أو زمن ضمن إطار معين يسمى جدولاً.

مثال:

الجدول التالي يوضح عدد الطلبة في بعض كليات المجتمع في عام 2001.

الكلية	عدد الطلبة
الكلية أ	300
الكلية ب	530
الكلية جـ	570
الكلية د	1200

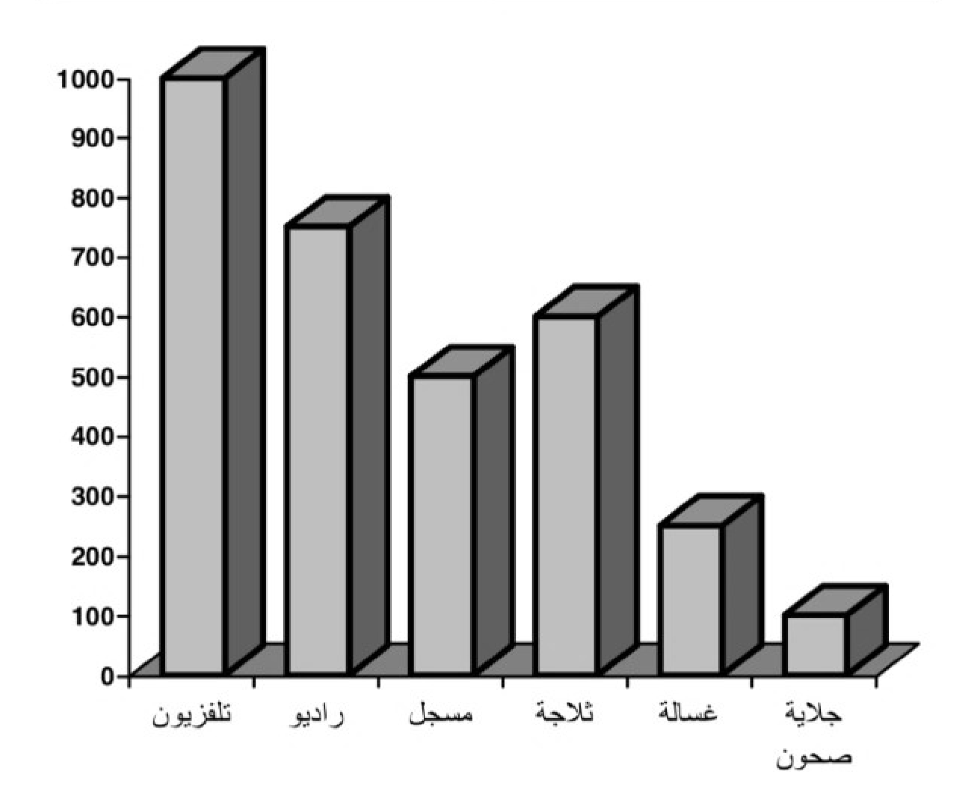
2- طريقة المستطيلات أو الأعمدة Rectangular method

يتم العرض بهذه الطريقة برسـم محـورين أفقـي وعمـودي يمثـل الأفقـي مـسمى الظاهرة (أو الزمن) والعمودي يمثل قيمـة الظـاهرة ثـم نرسـم مـستطيل قاعدتـه علـى المسمى، وطوله بقيمة المسمى، بمقياس رسم مناسب.

مثال:

الجدول التالي يمثل مبيعات شركة ما للأجهزة الكهربائية في سنة 2004 مثل هـذا الجدول بطريقة المستطيلات.

نوع الجهاز	عدد الأجهزة
تلفزيون	1000
راديو	750
مسجل	500
ثلاجة	600
غسالة	250
جلاية صحون	100



3- طريقة الخط البياني Line method

تستعمل لعرض تغير ظاهرة مع مسمى أو زمن وذلك برسم محورين أفقي وعمودي ورصد قيم الظاهرة مع الزمن أو المسمى بنقطة في المستوى على الصورة (المسمى أو الزمن، قيمة الظاهرة) ويكون التمثيل بنوعين من الخطوط:

أ- الخط المنكسر: ويكون بالوصل بين النقاط بخطوط مستقيمة.

ب- الخط المنحني: ويكون الوصل بين النقاط بخطوط منحنية.

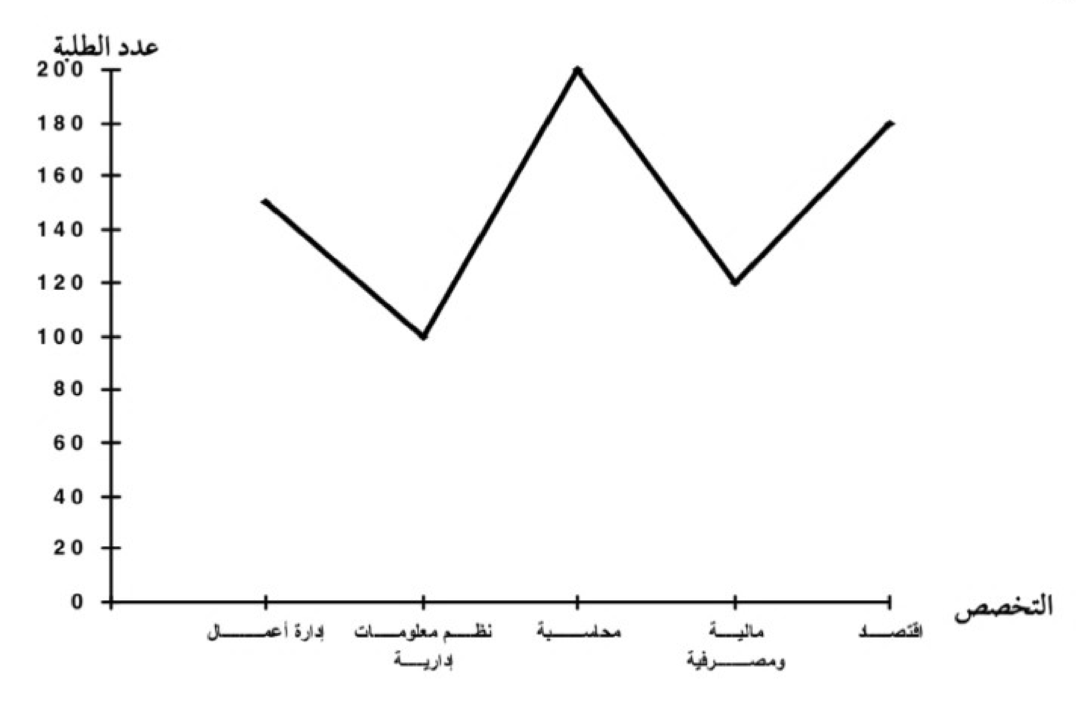
مثال:

الجدول التالي يمثل أعـداد الطلبـة في مـستوى البكـالوريوس في كليـة التجـارة في إحدى الجامعات الأردنية.

عدد الطلبة	التخصص
150	إدارة أعمال
100	نظم معلومات إدارية
200	محاسبة
120	مالية ومصرفية
180	اقتصاد

مثل هذا الجدول بطريقة الخط المنكسر:

الحل:



تمريـــــن:

مثل الجدول السابق بطريقة الخط المنحني.

4- طريقة الدائرة Circle method

وتكون بتقسيم الدائرة الكلية إلى قطاعات بنسب قيم الظاهرة ويحسب قياس زاوية القطاع بالطريقة التالية:

$$\text{قياس زاوية القطاع} = \frac{\text{قيمة الظاهرة}}{\text{المجموع الكلي لقيم الظاهرة}} \times 360°$$

مثال:

في سوق عمان المالي، إذا كان عدد الأسهم المباعة في أحد الأيام ممثلة بالجدول التالي:

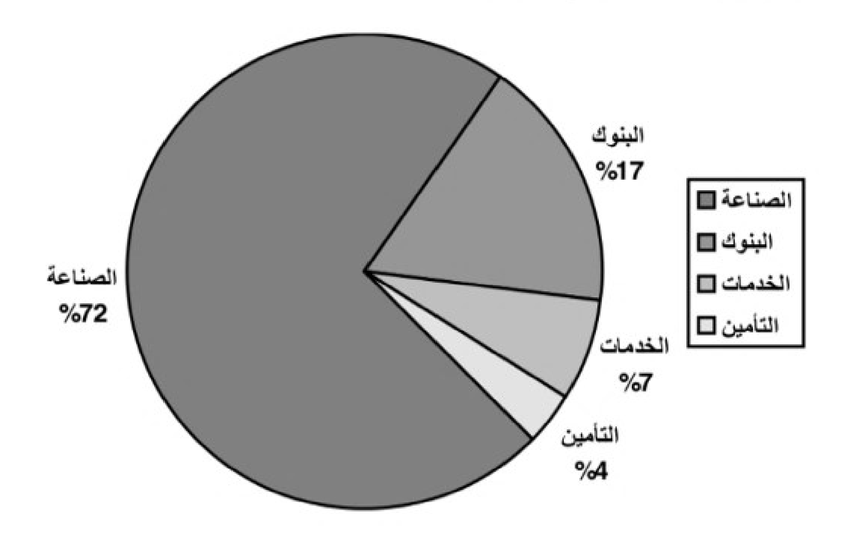
القطاع	عدد الأسهم
البنوك	29000
الخدمات	12000
التأمين	6000
الصناعة	123000
المجموع	170000

مثل هذا الجدول بطريقة الدائرة.

نجد أولاً قياس الزاوية لكل قطاع فيكون:

- زاوية قطاع البنوك $= \frac{29000}{170000} \times 360° \cong 61°$
- زاوية قطاع الخدمات $= \frac{12000}{170000} \times 360° \cong 25°$
- زاوية قطاع التأمين $= \frac{6000}{170000} \times 360° \cong 13°$
- زاوية قطاع الصناعة $= \frac{123000}{170000} \times 360° \cong 261°$

فيكون الجدول ممثلاً بالدائرة التالية:



5- طريقة الصور Picture method

في هذه الطريقة يتم تمثيل البيانات بصور لمجسم الظاهرة المراد دراستها، بشكل متناسب.

مثال:

في شركة لبيع البطاريات الجافة، كانت مبيعات الشركة لثلاث سنوات متتالية هي (10000) بطارية في عام 1990، (15000) بطارية في عام 1991، (17500) بطارية في عام 1992، اعرض هذه البيانات بطريقة الصور.

السنة	عدد البطارية
1990	🔋 🔋
1991	🔋 🔋 🔋
1992	🔋 🔋 🔋 🔋

مثلت كل (5000) بطارية بصورة بطارية واحدة.

التوزيعات التكرارية تمثيلها بيانياً Frequency Distribution

بناء جدول التوزيعات التكرارية: جدول التوزيع التكراري ما هو إلا وسيلة لاختصار حجم البيانات ووضعها في حيز مناسب يمكننا من الإحاطة بها من جميع جوانبها.

ولبناء جدول التوزيع التكراري نتبع الخطوات التالية والتي سندرجها ضمن المثال التالي:

مثال:

كون جدول التوزيع التكراري لعلامات (30) طالباً في امتحان ما:

46	49	48	58	54	50
40	62	37	48	54	75
54	48	59	45	34	58
47	61	49	44	68	39
63	56	43	57	41	45

1- نجد المدى المطلق (أو المدى) Range للبيانات: وهو

المدى المطلق = أكبر مشاهدة – أصغر مشاهدة

وفي هذه البيانات المدى المطلق = 75–34 =41

2- نحدد عدد فئات مبدئي مناسب لعدد البيانات:

وفي هذه البيانات نحدد عدد الفئات 7 فئات.

3- نجد طول الفئة: وهو

$$\text{طول الفئة} = \frac{\text{المدى المطلق}}{\text{عدد الفئات}}$$

وفي هذا التوزيع طول الفئة = $\frac{41}{7} = 5{,}7 \cong 6$

ملاحظة: إذا كان ناتج القسمة السابقة عدد غير صحيح نقربه لأقرب عدد صحيح.

4- نحدد الحدود الفعلية للفئات:

في البداية نجد الحد الأدنى الفعلي للفئة الأولى بأخذ أصغر قيمة في المشاهدات أو أقل منها بواحد ومن ثم نطرح منها نصف وحدة، ويكون في هذه المشاهدات هو

$$34 - 0.5 = 33.5$$

وبعدها نجد الحد الأعلى الفعلي للفئة وذلك بإضافة طول الفئة إلى الحد الأدنى الفعلي. فتصبح الفئة الأولى هي 33.5-39.5 ثم نحدد الفئات الأخرى بحيث يكون الحد الأدنى الفعلي للفئة هو الحد الأعلى الفعلي للفئة التي تسبقها. بحيث تكون آخر فئة تحوي أكبر مشاهدة.

وفي هذه البيانات تكون الحدود الفعلية للفئات هي:

33.5 – 39.5
39.5 – 45.5
45.5 – 51.5
51.5 – 57.5
57.5 – 63.5
63.5 – 69.5
69.5 – 75.5

وإذا أردنا إيجاد الفئات فإننا نضيف إلى الحد الأدنى الفعلي للفئة نصف وحدة ونطرح من الحد الأعلى الفعلي نصف وحدة فتكون الفئة الأولى مثلاً هي: 39-34.

ملاحظة: في بعض الحالات يكون عدد الفئات الناتج يختلف عن عدد الفئات المحدد في البداية وهذا ليس خطأ والسبب في ذلك هو كون البيانات هي بيانات منفصلة.

5- نفرغ البيانات في الجدول بوضع إشارة (/) لكل مشاهدة محتواه في الفئة وتكون الإشارة الخامسة مستعرضة وذلك لسهولة الجمع.

6- تجمع الإشارات لكل فئة لتكون تكرار الفئة.

7- نجد مركز الفئة والذي يكون:

مركز الفئة = $\frac{\text{الحد الأدنى + الحد الأعلى}}{2}$ = $\frac{\text{الحد الأدنى الفعلي+الحد الأعلى الفعلي}}{2}$

فمثلاً مركز الفئة الأولى = $36,5 = \frac{73}{2} = \frac{39 + 34}{2}$

بناءاً على الخطوات السابقة نكون جدول التوزيع التكراري التالي:

الفئات classes	الحدود الفعلية Boundaries	الإشارات	التكرار fi	مركز الفئة xi
34 – 39	33.5 – 39.5	///	3	36.5
40 – 45	39.5 – 45.5	//// /	6	42.5
46 – 51	45.5 – 51.5	//// ///	8	48.5
52 – 57	51.5 – 57.5	////	5	54.5
58 – 63	57.5 – 63.5	//// /	6	60.5
64 – 69	63.5 – 69.5	/	1	66.5
70 - 75	69.5 – 75.5	/	1	72.5

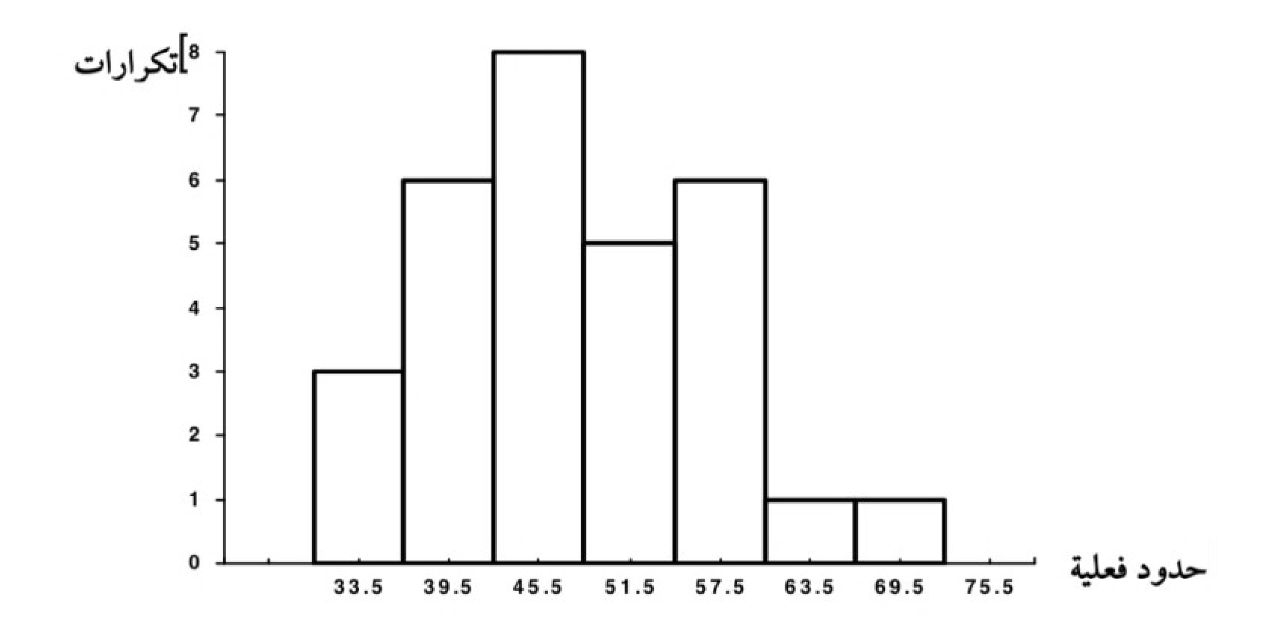
يسمى هذا الجدول والذي تكون أطوال فئاته متساوية توزيعاً تكرارياً منتظماً، أما إذا كانت فئاته غير متساوية الطول، فيدعى توزيعاً تكرارياً غير منتظم (انظر تمرين 12).

تمثيل التوزيعات التكرارية

1- المدرج التكراري:Frequency Histogram

وهو عبارة عن تمثيل كل فئة من فئات التوزيع بمستطيل حدود قاعدته الحدود الفعلية للفئات وارتفاعه يتناسب مع تكرار الفئة.

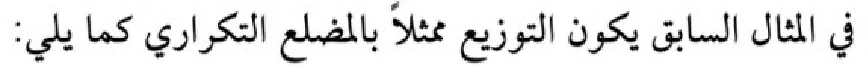
مثال: في المثال السابق، نمثل التوزيع بالمدرج التالي:

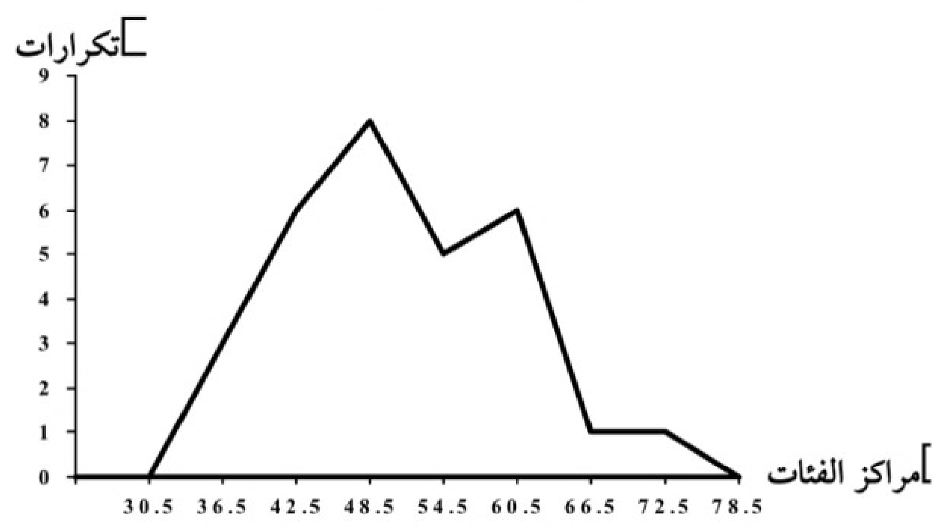


2- المضلع التكراري: Frequency Polygon

وهو مضلع مغلق نحصل عليه برسم محورين أفقي وعمودي ورصـد نقـاط مركـز الفئة على المحور الأفقي وتكرار الفئة على المحور العمودي لتكون الـنقط علـى الـصورة (مركز الفئة، التكرار) لتمثل رؤوس المضلع و نصل بين هذه النقاط بخطـوط مـستقيمة، ولإغلاق المضلع نأخذ مركز الفئة الـسابق للفئـة الأولى والـتي يكـون تكرارهـا صـفراً ومركز الفئة اللاحقة للفئة الأخيرة والتي يكون أيضاً تكرارها صفراً.

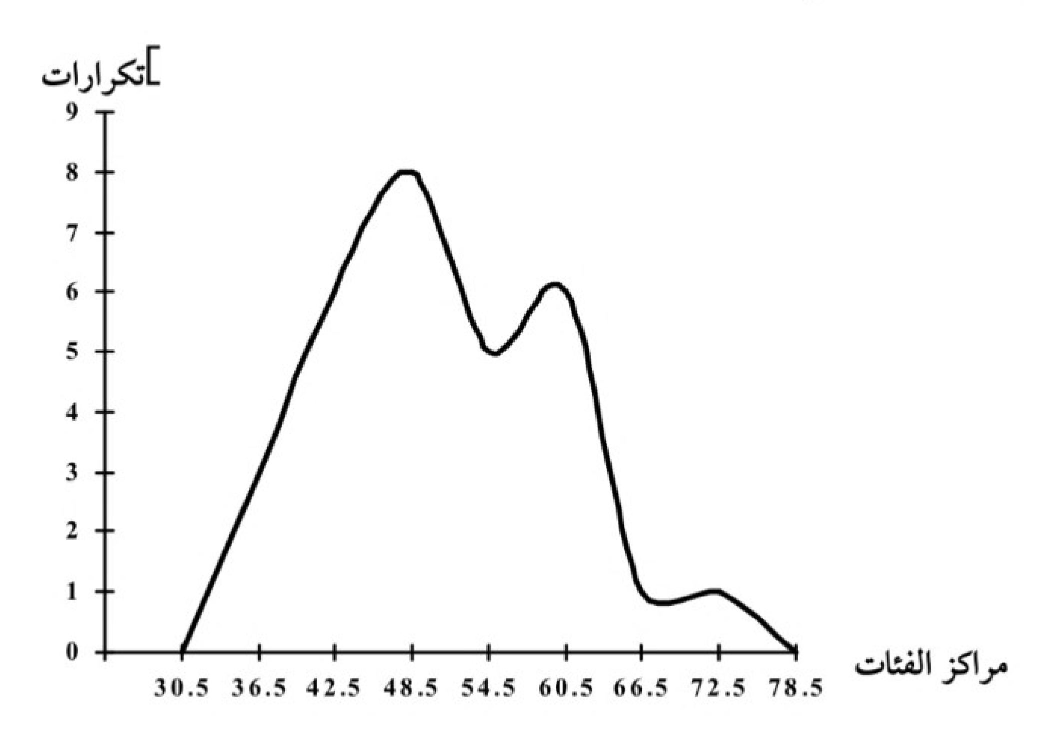
مثال:


في المثال السابق يكون التوزيع ممثلاً بالمضلع التكراري كما يلي:



3- المنحنى التكراري: Frequency Curve

إذا مهدنا المضلع التكراري وجعلناه منحنى بدلاً من خطوط مستقيمة فإننا نحصل على المنحنى التكراري.



ملاحظة: هنالك منحنى يشبه المنحنى التكراري ويرسم بنفس طريقة المنحنى التكراري ولكن النقط تكون الصورة (مركز الفئة، التكرار النسبي):

$$\text{حيث التكرار النسبي للفئة} = \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}}$$

4- المنحنى التكراري التراكمي (المنحنى المتجمع الصاعد)

Cumulative Frequency Curve

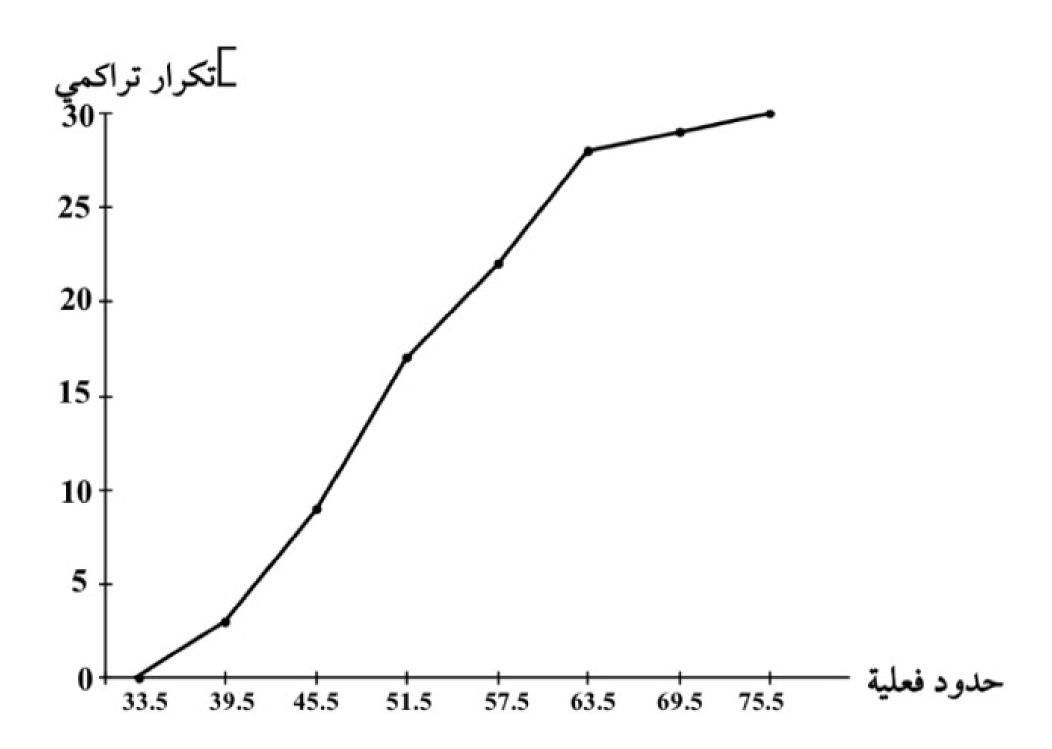
نحصل عليه برصد نقاط التكرار المتجمع على المحور العمودي مع الحد الأعلى الفعلي للفئات على المحور الأفقي.

والتكرار التراكمي للفئة: هو تكرار الفئة مضافاً إليه مجموع تكرارات الفئات التي تسبقها.

مثال:

في المثال السابق يكون المنحنى التكراري التراكمي للتوزيع هو المنحنى الناتج عـن رسم الجدول التالي:

الحدود الفعلية العليا Upper boundaries	التكرار التراكمي Cumulative Frequency
≤ 33.5	0
≤ 39.5	3
≤ 45.5	9
≤ 51.5	17
≤ 57.5	22
≤ 63.5	28
≤ 69.5	29
≤ 75.5	30



الجداول المقفلة والجداول المفتوحة Closed and open tables

تعريف:

1- يسمى الجدول الذي تكون فئته الأولى ليس لها حد أدنى جدولاً مفتوحاً من الأسفل.

2- ويسمى الجدول الذي تكون فئته الأخيرة ليس لها حد أعلى جدولاً مفتوحاً من الأعلى.

3- تسمى الجداول المفتوحة من أسفل ومن أعلى جداول مفتوحة.

4- تسمى الجداول غير المفتوحة من أسفل وغير المفتوحة من أعلى جداول مقفلة (كما في شرحنا السابق) وهذه الجداول الأكثر أهمية وشيوعاً.

مثال:

حدد نوع الجدول فيما يلي من حيث كونه مفتوحاً من أسفل، مفتوحاً من أعلى، مفتوحاً، أو مقفلاً.

أ-

الفئات Class	50-59	60-69	70-79	أكثر من 80
التكرار Frequency	4	5	2	7

ب-

الفئات Class	10-14	14-19	20-24	25-29
التكرار Frequency	1	5	7	2

جـ-

الفئات Class	أقل من 35	35-39	40-44	45-49
التكرار Frequency	9	2	5	4

د-

الفئات Class	أقل من 50	50-69	70-89	أكثر من 90
التكرار Frequency	12	5	4	19

الحل:

أ- مفتوح من أعلى. ب. مقفل. ج. مفتوح من أسفل. د.مفتوح.

تماريــن

1- إذا أجريت دراسة للبطالة على خريجي كليات المجتمع، فماذا يكون مجتمع الهـدف ومجتمع العينة في هذه الدراسة.

2- أعط مثال على كل نوع من أنواع العينات.

3- اختار عينة عشوائية بسيطة حجمها (20) من مجتمع إحصائي حجمه (2500).

4- الجـدول التـالي يمثـل أعـداد الخـريجين لكليـة الهندسـة في إحـدى الجامعـات في السنوات (2000-2004):

السنة	عدد الخريجين
2000	100
2001	125
2002	150
2003	150
2004	175

مثل الجدول بالطرق التالية:

أ- طريقة المستطيلات.

ب-طريقة الخط المنكسر.

ج- طريقة الدائرة.

د- طريقة الصور.

5- الجدول التالي يمثل مراكز الفئات وتكراراتها المقابلة:

xi	4	11	18	25	32	39	46
fi	3	6	10	15	9	5	2

أ- كون جدول التوزيع التكراري لهذا الجدول.

ب-مثل جدول التوزيع التكراري بالطرق التالية:

- المدرج التكراري.

- المضلع التكراري.
- المنحنى التكراري النسبي.
- المنحنى التكراري التراكمي.

6- المشاهدات التالية تمثل رواتب (40) موظفاً في دائرة حكومية:

117	200	137	145	113	117	115	110
225	230	113	145	115	225	250	113
185	180	175	113	200	113	250	117
137	148	248	237	245	240	195	190
219	213	209	173	167	195	194	188

كون جدول توزيع تكرار عدد فئاته عشرة فئات لهذه البيانات، ثم مثل هذه البيانات بالطرق التالية:

أ- المنحنى التكراري.

ب-المدرج التكراري.

ج- المنحنى التكراري التراكمي.

د- المنحنى التكراري النسبي.

7- قيست أقطار (20) كرة صغيرة بالسنتمتر مقربة لأقرب خانة عشرية واحدة، فكانت القياسات:

2.6	1.9	3.1	2.1	2.5
2.7	1.9	2.0	1.8	1.7
1.7	1.8	2.8	2.9	3.0
2.2	2.6	2.4	2.3	2.5

كون جدول توزيع تكراري عدد فئاته(5) ومثله بمدرج تكراري

(إرشاد: الحد الفعلي الأدنى للفئة = الحد الأدنى -0.05).

8- جدول توزيع تكراري طول فئة (9) أخذت منه فئة مركزها (38) أوجد الحدود الفعلية لتلك الفئة.

9- فئة من توزيع تكراري مركزها 16 وحدها الأدنى الفعلي (12.5) جد طول هذه الفئة.

10- ثلاثة مصانع A, B, C للألبسة إذا كان مجموع إنتاجها في عام 1997 هو (10000) قطعة، إذا مثل إنتاج هذه المصانع باستخدام طريقة القطاعات الدائرية، وكانت زاوية قطاع إنتاج المصنع (B) هي (162°)، فما هو حجم إنتاج هذا المصنع في ذلك العام.

11- فئة تكرارها (8) وتكرارها النسبي 0.32 جد مجموع تكرارات جدول التوزيع التكراري الذي أخذت منه هذه الفئة.

12- مثل جدول التوزيع غير المنتظم التالي بمدرج تكراري.

الفئات Class	10-19	20-29	30-49	50-64	د
التكرار Frequency	5	10	16	14	5

13- مركز الفئة الثانية =

14- مركز الفئة الخامسة =

15- طول الفئة =



الوحدة الثانية

مقاييس النزعة المركزية

Measures of Central Tendency

مقاييس النزعة المركزية

Measures of Central Tendency

مقدمة

مقياس النزعة المركزيةMeasure of Central Tendency هو تلك القيمة التي نجدها من مجموعة البيانات (Data) التي لدينا والتي تمثل هذه البيانات بشكل مقبول.

وهنالك عدة مقاييس للنزعة المركزية منها الوسط الحسابي Mean والوسيط Median والمنوال Mode. ونفضل واحد منها على الآخر حسب البيانات التي لدينا.

وقبل أن نبدأ بطرق إيجاد تلك المقاييس، سنتعرف على موضوع المئينات Percentiles والرتب المئينية والعشيرات Deciles والربيعات Quartiles أولا وذلك من أجل التسلسل في العرض.

المئينات Percentiles

تعريف: المئين k هو تلك المشاهدة التي يقل عنها أو يساويها k% من المشاهدات. وسنرمز للمئين k بالرمز P_k.

فمثلاً: P_{60} هو المشاهدة التي يقل عنها أو يساويها 60٪ من المشاهدات، وطبعا يزيد عنها 40٪ من المشاهدات.

مثال:

ما هي المشاهدة التي يزيد عنها $\frac{1}{4}$ المشاهدات.

الحل:

المشاهدة التي يزيد عنها $\frac{1}{4}$المشاهدات هي تلك المشاهدة التي يقل عنها أو يساويها $\frac{3}{4}$المشاهدات. أي يقل عنها أو يساويها 75٪ من المشاهدات أي تلك المشاهدة هيP_{75}.

ولإيجاد المئينات Percentiles للجـداول التكراريـة سـنتبع خطـوات موضـحة في المثال التالي.

مثال: إذا كان لدينا التوزيع التالي:

Class	10-14	15-19	20-24	25-29	30-34	Total
Frequency	1	5	8	4	2	20

أوجد: P_{50}.

الحل:

أولاً: نكتب الجدول التكراري التراكمي للتوزيع.

Upper boundaries	**Cumulative Frequency**
Less than or equal 14.5	1
Less than or equal 19.5	6
Less than or equal 24.5	14
Less than or equal 29.5	18
Less than or equal 34.5	20

ثانياً: التكرار التراكمي Cumulative Frequency للمئين 50:

$$C.F(P_{50}) = \frac{50}{100} \times Total\ \ frequency$$

$$C.F(P_{50}) = \frac{50}{100} \times 20$$

$$= 10$$

وبالنظر للجدول التكراري التراكمي نجد أن 10 تقـع بـين التكـرارين التراكمـيين 6، 14 فيكون P_{50} واقعاً بين الحدين الفعليين 19.5 ، 24.5.

$$\begin{array}{lll} & 19.5 \ldots\ldots\ldots\ldots 6 & \\ 5 & P_{50} \ldots\ldots\ldots\ldots 10 \quad 4 & 8 \\ & 24.5 \ldots\ldots\ldots\ldots 14 & \end{array}$$

ثالثاً: بالنسبة والتناسب نجد قيمة P_{50}

$$P_{50} = 19.5 + \frac{4}{8} \times 5$$

$$= 22$$

ملاحظة: تسمى الفئة التي تحوي المئين k بالفئة المئينية لـذلك المـئين. ففـي مثالنـا السابق تكون الفئة المئينية للمئين 50 هي 19.5 – 24.5

Class	10-19	20-29	30-39	40-49	50-59	Total
Frequency	7	13	30	19	11	80

أوجد: 1) P_{25} 2) P_{60} 3) P_{75}

الحل:

نحول الجدول إلى جدول تكراري تراكمي.

Upper Boundaries	**Cumulative Frequency**
Less than 19.5	7
Less than 29.5	20
Less than 39.5	50
Less than 49.5	69
Less than 59.5	80

1- نجد الترتيب التراكمي للمئين 25:

$$C.F\left(P_{25}\right) = \frac{25}{100} \times 80 = 20$$

وبالنظر للجدول التكراري التراكمي نرى أن هذه القيمة تقابل 29.5.

$$\therefore\ P_{25} = 29.5 \square$$

2- نجد الترتيب التراكمي للمئين 60:

$$C.F\left(P_{60}\right) = \frac{60}{100} \times 80 = 48$$

وبالنظر للجدول التكـراري التراكمـي نجـد أن هـذه القيمـة تقـع بـين التكـرارين التراكميين 20، 50.

إذن P_{60} واقعاً بين 29.5 ، 39.5.

وبالنسبة والتناسب نجد P_{60}.

10	29.5 20	28	30
	P_{60} 48		
	39.5 50		

$$P_{60} = 29.5 + \frac{28}{30} \times 10$$
$$= 29.5 + 9.3$$
$$= 38.8$$

3- نجد الترتيب التراكمي للمئين 75 :

$$C.F\left(P_{75}\right) = \frac{75}{100} \times 80 = 60$$

وبالنظر للجدول التكراري التراكمي نجد أن هذه القيمة واقعة بين التكرارين التراكميين 50، 69.

10	39.5 50	10	19
	P_{75} 60		
	49.5 69		

$$P_{75} = 39.5 + \frac{10}{19} \times 10$$
$$= 39.5 + 5.26$$
$$= 44.76$$

تعريف: الرتبة المئينية Percentile Rank لمشاهدة ما هي النسبة المئوية للتكرار التراكمي المقابل لتلك المشاهدة بالنسبة إلى مجموع التكرارات.

ولإيجاد الرتبة المئينية نتبع خطوات نوضحها في المثال التالي:

مثال:

للجدول التكراري Frequency Distribution التالي

Class	20-24	25-29	30-34	35-39	40-44	Total
Frequency	7	9	20	8	6	50

أوجد الرتبة المئينية

a- للمشاهدة 32   b- للمشاهدة 21

الحل:

a-

أولاً: نحول الجدول إلى جدول تكراري تراكمي

Upper Boundaries	**Cumulative Frequency**
Less than 24.5	7
Less than 29.5	16
Less than 34.5	36
Less than 39.5	44
Less than 44.5	50

ثانياً: نبحث عن موقع المشاهدة (Observation) 32 في الجدول التكراري التراكمي، وضمن الحدود الفعلية للفئات وليس ضمن التكرارات التراكمية.

فنجد أن هذه القيمة واقعة بين الحدين الفعليين 29.5 ، 34.5.

ثالثاً: وبطريقة النسبة والتناسب نجد التكرار التراكمي المقابل لتلك المشاهدة.

29.5............................16

32 ....C.F(32)

24.5...........................36

5   2.5   20

حيث التكرار التراكمي للمشاهدة (32)

$$C.F(32) = 16 + \frac{2.5}{5} \times 20$$
$$= 16 + 10$$
$$= 26$$

رابعاً: تكون الرتبة المئينية للمشاهدة 32: P.R(32)

$$P.R(32) = \frac{C.F(32)}{\text{Total of frequency}} \times 100\%$$

$$= \frac{26}{50} \times 100\%$$

$$= 52\%$$

b- نلاحظ هنا أن المشاهدة (21) أقـل (24.5) لـذلك نـضيف فئـة سـابقة حـدها الفعلـي الأعلى (19.5) ويكون التكرار التراكمي الذي أقل من أو يساوي 19.5 هو (Zero).

5	1.5	19.5..........................Zero	7
		21 ....C.F(21)	
		24.5..........................7	

$$\therefore C.F(21) = 7 \times \frac{1.5}{5} + 0$$

$$= 2.1$$

$$\therefore P.R(21) = \frac{2.1}{50} 100\%$$

$$= 4.2\%$$

مثال:

إذا كانت رواتب (60) عاملاً في مصنع ما موزعة كما في الجدول التالي:

فئات الرواتب بالدينار	80-89	90-99	100-109	110-119	120-129	Total
التكرارات	6	14	20	13	7	60

أوجد:

1. الرتبة المئينية للراتب 95.
2. الرتبة المئينية للراتب 109.5.
3. الرتبة المئينية للراتب 117.

الحل:

نحول الجدول إلى جدول تكراري تراكمي

Upper Boundary	Cumulative Frequency
Less then 89.5	6
Less then 99.5	20
Less then 109.5	40
Less then 119.5	53
Less then 129.5	60

1. الراتب 95 يقع بين الحدين الفعليين 89.5 ، 99.5.

الآن بطريقة النسبة والتناسب نجد التكرار التراكمي المقابل للراتب 95

10	5.5	89.5..........................6		14
		95 ....C.F(95)		
		99.5..........................20		

$$C.F(95) = \frac{5.5}{10} \times 14 + 6$$

$$= 7.7 + 6$$

$$= 13.7$$

$$P.R(95) = \frac{13.7}{60} \times 100\%$$

$$= 22.83\%$$

لاحظ أن النسبة 22.83٪ تعني أن 22.83 من العمال رواتبهم تقل عن أو تساوي 95 دينار.

2. ننظر للجدول التكراري التراكمي فنجد أن الراتب 109.5 يقابل تكرار تراكمي مقداره 40 لذا فإن

$$P.R\ (109.5) = \frac{40}{60} \times 100\%$$
$$= 66.67\%$$

- **سؤال:** ماذا نعني بالنسبة 66.67%؟

3. بالنظر للجدول التكراري التراكمي نجد أن الراتب 117 يقع بين الحدين الفعلـيين 109.5 ، 119.5 وبالنسبة والتناسب.

10	7.5	109.5	40	13	
		117	C.F(117)		
		119.5	53		

نجد أن التكرار التراكمي المقابل للراتب 117 =

$$C.F\ (117) = 40 + \frac{7.5}{10} \times 13$$
$$= 40 + 9.75$$
$$= 49.57$$
$$P.R(117) = \frac{49.75}{60} \times 100\%$$
$$= 82.92\%$$

- **سؤال:** ماذا نعني بالنسبة 82.92%.

إيجاد المئينات بيانياً Finding Percentiles by using graphs

لتوضيح عملية إيجاد المئينات بيانياً لنأخذ المثال التالي:

مثال:

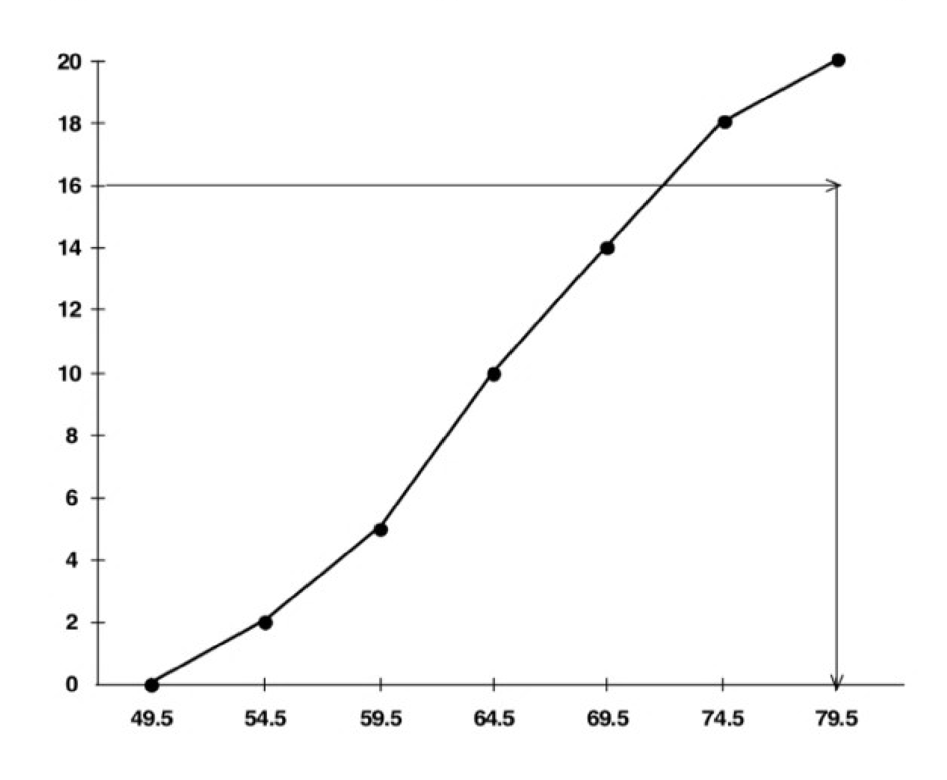
أوجد P_{80} للجدول التكراري التالي بيانياً.

Class	50-54	55-59	60-64	65-69	70-74	75-79	Total
Frequency	2	3	5	4	4	2	20

الحل:

أولاً: نرسم المنحنى التكراري التراكمي. بعد كتابة الجدول التكراري التراكمي:

Upper boundary	**Cumulative Frequency**
Less then 54.5	2
Less then 59.5	5
Less then 64.5	10
Less then 69.5	14
Less then 74.5	18
Less then 79.5	20



لاحظ: عند رسم المنحنى التكراري التراكمي يكـون التكـرار التراكمـي المقابـل للحـد الفعلي 49.5 مساوياً 0.

ثانياً:

$$C.F(80) = \frac{80}{16} \times 20$$

$$= 16$$

ثالثاً: على المحـور الرأسـي Vertical axis وعنـد القيمـة 16 نقـيم عمـود، فيقطـع المنحنـى التكراري التراكمي في نقطة ما، ومن تلك النقطة نسقط عمود على المحـور الأفقـي، فيقطـع المحـور الأفقي Horizontal axis عند القيمة 72 فتكون هذه القيمة هي المئين 80.

$\therefore P_{80}=72.$

لاحظ أن عملية إيجاد المئين بهذه الطريقـة تعتمـد علـى دقـة الـشخص في عمليـة الرسم لذلك فالجواب يكون تقريبياً.

العشيرات والربيعات Deciles and Quartiles

تعريف: العشير (decile) k هـو المـشاهدة الـتي تقـل عنهـا أو يـساويها $\frac{k}{10}$ مـن مجموع التكرارات. وسنرمز للعشير k بالرمز D_k.

فمثلاً العشير الثـاني هـو المـشاهدة الـتي يقـل عنهـا أو يـساويها $\frac{2}{10}$ مـن مجمـوع التكرارات وهو نفس المئين P_{20}.(أي $D_2=P_{20}$)

كذلك العشير السادس هو المشاهدة التي يقـل عنهـا أو يـساويها $\frac{6}{10}$ مـن مجمـوع التكرارات وهو نفس المئين P_{60}. (أي $D_6=P_{60}$)

إذن فيمكننا القول أن العشير k هو المئين 10K. ($D_k=P_{10k}$)

تعريف (Definition):

1- **الربيع الأول (First Quartile):** هو المـشاهدة الـتي يقـل عنهـا أو يـساويها $\frac{1}{4}$ مجمـوع التكـرارات وسـنرمز لـه بـالرمز Q_1. ويـسمى كـذلك بـالربيع الأدنـى. (Lower Quartile)

2- **الربيع الثاني(Second Quartile)**: هو المشاهدة التي يقل عنها أو يساويها $\frac{2}{4}$ مجموع التكرارات وسنرمز له بالرمز Q_2. ويسمى كذلك بالربيع الأوسط.Midle Quartile

3- **الربيع الثالث (Third Quartile)**: هو المشاهدة التي يقل عنها أو يساويها $\frac{3}{4}$ مجموع التكرارات وسنرمز له بالرمز Q_3. ويسمى كذلك الربيع الأعلى (Upper Quartile).

لاحظ أن المشاهدة التي يقل عنها $\frac{1}{4}$ مجموع التكرارات هي نفس المشاهدة التي يقل عنها 25٪ من مجموع التكرارات أي أن:

$Q_1=P_{25}$ وكذلك $Q_2=P_{50}$، وأيضاً $Q_3=P_{75}$

وبعد التعرف على العشيرات والربيعات يمكننا القول أنه لحساب العشيرات والربيعات نستخدم الطريقة المتبعة في حساب المئينات.

مقاييس النزعة المركزية Measures of Central Tendency

أولاً: الوسط الحسابي Arithmatic Mean

أ- في حالة المشاهدات المفردة:

تعريف: إذا كان لدينا المفردات $x_1 , x_2, \ldots , x_n$ فيعرف الوسط الحسابي $(\bar{x})$ لهذه المفردات بالعلاقة

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

حيث n عدد المفردات

أو باستخدام رمز المجموع

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

وبشكل مختصر أكثر

$$\bar{x} = \frac{\sum x}{n}$$

مثال:

أوجد الوسط الحسابي للمفردات 3، 7، 6، 5، 9؟

الحل:

$$\bar{x} = \frac{3+7+6+5+9}{5}$$
$$\bar{x} = \frac{30}{5}$$
$$\bar{x} = 6$$

مثال:

إذا كان مجموع ما مع سبعة طلاب (105) دنانير، فما هو الوسط الحسابي لما ما مع هؤلاء الطلبة من الدنانير؟

$$\sum x = 105 \qquad , \qquad n = 7$$
$$\therefore \ \bar{x} = \frac{105}{7} = 15$$

مثال:

إذا كان الوسط الحسابي (Mean) لعلامات عدد من الطلاب هو (63)، وكان مجموع علاماتهم (1071). فما عدد هؤلاء الطلبة؟

الحل:

نفرض أن عدد هؤلاء الطلبة n فيكون

$$\bar{x} = \frac{\sum x}{n}$$
$$\Rightarrow 63 = \frac{1071}{n}$$
$$\Rightarrow n = 17$$

ب- في حالة المشاهدات المتكررة: (Weighted mean) الوسط الموزون

تعريف: إذا كان لدينا $x_1, x_2, \ldots, x_n$ مجموعة من المشاهدات وكانت تكرارات هذه المشاهدات $f_1, f_2, \ldots, f_n$ فيعرف الوسط الحسابي لها بالعلاقة

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

أو بشكل مختصر $\bar{x} = \frac{\sum fx}{\sum f}$ ويسمى هذا الوسط بالوسط المرجح (الموزون).

مثال:

إذا كانت علامات طالب في (10) مواد موزعة كما في الجدول التالي

العلامة (x)	60	70	85	90	Total
عدد المواد (f)	2	3	4	1	10

احسب الوسط الحسابي لعلامات هذا الطالب؟

الحل:

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$\bar{x} = \frac{60 \times 2 + 70 \times 3 + 85 \times 4 + 90 \times 1}{10}$$

$$\bar{x} = \frac{760}{10}$$

$$\bar{x} = 76$$

مثال:

إذا كان معدل رواتب عمال ثلاثة مصانع معطاة كالآتي:

	معدل الرواتب	عدد العمال
المصنع أ	150 دينار	60
المصنع ب	140 دينار	50
المصنع جـ	100 دينار	90

جد معدل رواتب العمال في المصانع الثلاثة معا؟

الحل:

$$\bar{x} \text{ (معدل الرواتب) } = \frac{90\times100+50\times140+60\times150}{90+50+60}$$
$$= 125 \text{ J.D.}$$

جـ- في حالة التوزيعات التكرارية Frequency distributions

في هذه الحالة سنجد الوسط الحسابي بطريقتين وهما

1- الطريقة العامة لإيجاد الوسط الحسابي

تعريف: إذا كان لدينا جدول تكراري فيه m من الفئات مراكزها (Class-marks) هي $x_1, x_2, \ldots, x_m$ وتكراراتها المقابلة $f_1, f_2, \ldots, f_m$ على الترتيب فنعرف الوسط الحسابي.

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i x_i}{\sum_{i=1}^{m} f_i}$$

أو بشكل مختصر $\bar{x} = \frac{\sum fx}{\sum f}$ حيث x هي مركز الفئة

مثال:

أوجد الوسط الحسابي للجدول التكراري التالي:

Class	20-24	25-29	30-34	35-39	40-44	Total
Frequency	9	17	20	9	5	60

الحل:

$$\bar{x} = \frac{\sum fx}{\sum f}$$

حيث x: مركز الفئة

f: تكرار الفئة

class	f	Class mark(x)	xf
20-24	9	22	198
25-29	17	27	459
30-34	20	32	640
35-39	9	37	333
40-44	5	42	210
Total	60		1840

$$\therefore \bar{x} = \frac{1840}{60}$$

$$\bar{x} = 30.67$$

2- طريقة الوسط الفرضي Assumed mean

لو أخذنا الأعداد 21، 22، 23، 24، 25 وطلبنا منك أن تجد الوسط الحسابي لها. فيمكنك إيجاده بطريقة غير مباشرة وهي أن تطرح من كل عدد من هذه الأعداد (20) فتصبح الأعداد هي 1، 2، 3، 4، 5 على الترتيب ثم تجد لهذه الأرقام الوسط الحسابي وهو $\frac{1+2+3+4+5}{5} = 3$ ثم تضيف إلى هذه القيمة (20) فتصبح 23 وهو الوسط الحسابي للأعداد المطلوبة.

هذه الطريقة في إيجاد الوسط الحسابي تسمى طريقة الوسط الفرضي.

ويكون قانون الوسط الحسابي بطريقة الوسط الفرضي (A) للمشاهدات المفردة $x_1, x_2, \ldots, x_n$ هو.

$$\bar{x} = A + \frac{\sum di}{n}$$

حيث $d_i = x_i - A$

تمرين:

في المناقشة السابقة اطرح من كل مشاهدة العدد 15 (A=15) واتبع نفس الأسلوب المستخدم لتجد الوسط الحسابي للأعداد المطلوبة؟ ولاحظ هل يتغير الجواب؟

والآن لنكتب القانون المستخدم في حساب الوسط الحسابي للجداول التكرارية بطريقة الوسط الفرضي.

$$\bar{x} = A + \frac{\sum fd}{\sum f}$$

حيث: A الوسط الفرضي

d= x-A

أي أن d هي انحراف مركز الفئة عن الوسط الفرضي.

وبالنسبة لكيفية اختيار A فلا يوجد عليها أية قيود ولكن لتسهيل العمليات الحسابية نختار A مركز إحدى الفئات ذات الموقع المتوسط وتكرارها كبيراً نوعاً ما.

مثال:

أوجد الوسط الحسابي للجدول التكراري التالي بطريقة الوسط الفرضي.

Class	20-29	30-39	40-49	50-59	60-69	70-79	Total
Frequency	3	4	6	7	6	4	30

الحل:

$$\bar{x} = A + \frac{\sum fd}{\sum f}$$

الآن نكتب الجدول التكراري بعد إيجاد مراكز فئاته. ولتكن A=54.5

Class	f	D= X-A	fd
20-29	3	-30	-90
30-39	4	-20	-80
40-49	6	-10	-60
50-59	7	0	0
60-69	6	10	60
70-79	4	20	80
Total	30		-90

$$\bar{x} = 54.4 + \frac{-90}{30}$$
$$= 54.5 - 3$$
$$= 51.5$$

ملاحظة: لا تتغير قيمة الوسط الحسابي بتغير الوسط الفرضي.

تمريـــن:

احسب الوسط الحسابي للجدول التكراري في المثال السابق بالطريقة العامة؟

أهم خصائص الوسط الحسابي

1- الوسط الحسابي يتأثر بالعمليات الحسابية الأربع: فمثلاً إذا كان لدينا المفردات $x_1, x_2, \ldots, x_n$ وسطها الحسابي $\bar{x}$ وعدلت هذه المفردات حسب العلاقة $y= ax+b$ ، حيث a , b عددان حقيقيان، x المفردة قبل التعديل، y المفردة بعد التعديل، فإن $\bar{y} = a\bar{x} + b$ ، حيث $\bar{y}$ الوسط الحسابي للمفردات بعد التعديل.

مثال:

إذا كان الوسط الحسابي لمجموعة من المفردات 60 وعدّلت جميع المفردات حسب العلاقة $y=200-2x$. فأوجد الوسط الحسابي بعد التعديل؟

الحل: ليكن

الوسط الحسابي بعد التعديل = $\bar{y}$

الوسط الحسابي قبل التعديل = $\bar{x}$

فيكون $\bar{y} = 200 - 2\bar{x}$

$$\bar{y} = 200 - 2\,(60)$$
$$= 200 - 120$$
$$= 80$$ (الوسط بعد التعديل)

2- مجموع انحرافات القيم عن وسطها الحسابي يساوي (Zero).

فمــثلاً إذا كــان لــدينا المفــردات $x_1, \quad x_2, \quad \ldots, \quad x_n$ وســطها الحــسابي $\bar{x}$ فإن $\sum(x-\bar{x}) = 0$

الحل:

$$\bar{x} = \frac{1+4+7+5+3}{5}$$
$$= \frac{20}{5} = 4$$
$$\sum(x-\bar{x}) = (1-4)+(4-4)+(7-4)+(5-4)+(3-4)$$
$$= 0$$

مثال:

إذا كانت انحرافات 3 قيم عن وسطها الحسابي هي a , 3 , 2 فأوجد قيمة a؟

الحل:

$2+3+a = 0 \Rightarrow a = -5$

3- الوسط الحسابي يتأثر بالقيم الشاذة "المتطرفة"

فمثلاً إذا كان لدينا ثلاثة أشخاص معهم بالـدينار، 2، 4، 900 فيكـون الوسـط الحسابي لما ما مع هؤلاء الأشخاص هـو $302 = \frac{2+4+900}{3}$ دينـار وهـذا لا يعطـي صورة واقعية لما ما مع هـؤلاء الأشـخاص. لـذلك نلجـأ إلى مقـاييس أخـرى لا تتـأثر بالقيم المتطرفة مثل الوسيط وغيره.

ثانياً: الوسيط Median

الوسيط (Me) هو مقياس آخر من مقاييس النزعـة المركزيـة وهـو المـشاهدة الـتي يكون مجموع التكرارات التي تسبقها مساوياً لمجموع التكرارات التي تليها وبلغة المئينـات يكون الوسيط هو المئين خمسين.

أي أن الوسيط = المئين 50 = العشير 5= الربيع الثاني.

وإذا رمزنا للوسيط بالرمز Me فإن

$Me = P_{50} = D_5 = Q_2$

والآن لنأتي لطرق إيجاد الوسيط.

أ- في حالة المفردات

لإيجاد الوسيط في حالة المفردات والتي عددها n، نرتب هذه المشاهدات تـصاعدياً ويكون :

الوسيط = المشاهدة التي ترتيبها $\frac{n+1}{2}$ إذا كان n فردياً
الوسيط الحسابي للمشاهدتين اللتين ترتيبهما $\frac{n}{2}$، $\frac{n}{2}+1$ زوجيا

مثال:

أوجد الوسيط للمفردات 18 ,10 ,7 ,16 ,9 ,7 ,1

الحل:

نرتب المشاهدات تصاعدياً كما يلي: 18 ,16 ,10 ,9 ,7 ,7 ,1

لاحظ أن عدد المشاهدات n=7، أي أن الوسيط هو المشاهدة التي ترتيبها

$$\frac{7+1}{2}=4$$

إذن فالوسيط هو المشاهدة الرابعة وهي Me=9.

مثال:

أوجد الوسيط للمفردات: 20 ,16 ,12 ,9 ,6 ,5 ,4

الحل: المشاهدات مرتبة تصاعدياً

n=8 ⇐ الوسيط هو الوسط الحسابي للمشاهدتين $\frac{8}{2}+1$, $\frac{8}{2}$

أي أن الوسيط هو الوسط الحسابي للمشاهدتين الرابعة والخامسة

$$Me=\frac{9+12}{2}=10.5$$

ب- في حالة الجداول التكرارية

لحساب الوسيط في الجداول التكرارية نحسب المئين 50

مثال: أوجد الوسيط للجدول التكراري التالي:

Class	10-14	15-19	20-24	25-29	Total
Frequency	4	8	5	3	20

الحل:

Upper Boundary	Cumulative Frequency
Less than 14.5	4
Less than 19.5	12
Less than 24.5	17
Less than 29.5	20

$$C.F(Me) = \frac{50}{100} \times 20$$
$$= 10$$

تقع بين 4، 12

5	14.5 4	6	8
	Me=P50 10		
	19.5 12		

$$Me = 14.5 + \frac{6}{8} \times 5$$
$$= 14.5 + 3.75$$
$$= 18.25$$

- **سؤال** : أوجد الوسيط في المثال السابق بيانياً.

تعريف : الفئة التي تحوي الوسيط تسمى الفئة الوسيطية.

- **سؤال** : في المثال السابق أوجد الفئة الوسيطية؟

ثالثاً: المنوال Mode

المنوال (Mo) هو المقياس الثالث من مقاييس النزعة المركزية وهو المفردة (المشاهدة) الأكثر تكراراً.

أ- إيجاد المنوال في حالة المفردات

إذا كان لدينا مجموعة من المفردات فيكون منوالها هو المفردة الأكثر تكرارا.

مثال:

أوجد المنوال للمفردات 3 ,2 ,4 ,5 ,9 ,7 ,2

الحل:

$Mo = 2$

إذا وجد أكثر من مشاهدة لها نفس التكرار، ويزيد عن باقي تكرارات المشاهدات الأخرى فتكون هذه المشاهدات منوالات.

مثال:

أوجد المنوال للمفردات 9 ,2 ,5 ,7 ,10 ,2 ,9 ,7 ,1

الحل: يوجد ثلاثة منوالات هي 9 ,7 ,2

إذا كانت جميع المشاهدات لها نفس العدد من التكرارات فنقول أنه لا يوجد منوال.

مثال:

أوجد المنوال للمفردات 7 , 16 , 8 ,10 , 2

الحل:

لا يوجد منوال.

مثال:

أوجد المنوال للمفردات 7 ,9 ,5 ,5 ,9 ,7 ,9 ,5 ,7

الحل:

لا يوجد منوال.

ب- إيجاد المنوال للجداول التكرارية

المنوال للجداول التكرارية هو مركز الفئة التي تقابل أكبر تكرار، وتلك الفئة تسمى الفئة المنوالية.

مثال:

أوجد المنوال للجدول التكراري التالي:

Class	30-34	35-39	40-44	45-49	50-54	Total
Frequency	10	10	15	8	7	50

الحل:

الفئة المنوالية هي 40-44 فيكون

$$Mo = \frac{40 + 44}{2}$$
$$= 42$$

مثال:

أوجد المنوال للجدول التكراري التالي:

Class	2-4	5-7	8-10	11-13	14-16	Total
Frequency	2	9	3	9	7	30

الحل:

يوجد منوالان هما $\frac{11 + 13}{2}$, $\frac{5 + 7}{2}$

أي أن المنوالان هما 12 , 6

تعريف: التوزيعات التي لها منوال واحد تسمى أحادية المنوال، والتي لها منوالان تسمى ثنائية المنوال والتي لها أكثر من منوالين تسمى عديدة المنوالات. أو متعددة المنوالات.

العلاقة الخطية بين الوسط والوسيط والمنوال

في التوزيعات أحادية المنوال لوحظ أن هنالك علاقة خطية تربط مقاييس النزعة المركزية. مع التأكيد أن هذه العلاقة مبينة على التجربة والملاحظة. أي أن هذه العلاقة ليست دقيقة ولكنها تقريبية وهذه العلاقة هي:

(Mean-Mode) = 3(Mean-Median)

وبالرموز $\left(\bar{X} - Mo\right) = 3\left(\bar{X} - Me\right)$

أي أن بعد الوسط الحسابي عن المنوال ثلاثة أمثال بعده عن الوسيط.

مثال:

إذا كان الوسط الحسابي لتوزيع أحادي المنـوال يـساوي 60 والمنـوال 50 فأوجـد تقدير الوسيط لهذا التوزيع؟

$$\left(\overline{X} - Mo\right) = 3\left(\overline{X} - Me\right)$$

$$(60 - 50) = 3(60 - Me)$$

$$10 = 180 - 3\, Me$$

$$3Me = 170$$

$$\Rightarrow Me = \frac{170}{3}$$

$$= 56.7$$

العزوم والالتواء والتفرطح Moments, Skewness and Kurtosis

تعريف: إذا كان لدينا المشاهدات $x_1, x_2, \ldots, x_n$

وكان a عدداً حقيقياً فإن العزم الرائي (r^{th} moment)حول العدد a

$$mr(a) = \frac{\sum_{i=1}^{n}\left(x_i - a\right)^r}{n} \qquad r \in \{1,2,3,\ldots\}$$

العزم الرائي حول نقطة الأصل (the origin)

$$mr(o) = \frac{\sum_{i=1}^{n}\left(x_i\right)^r}{n}$$

العزم الرائي حول الوسط الحسابي $\left(\overline{X}\right)$ هو:

$$mr(\overline{X}) = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^r}{n}$$

تعريف:

إذا كان لدينا جـدول تكـراري مراكـز فئاتـه $x_1, x_2, \ldots, x_n$ والتكـرارات المقابلـة لتلك الفئات $f_1, f_2, \ldots, f_n$ على الترتيـب فـإن العـزم الرائـي حـول العـدد الحقيقـي a يعرف بالقانون:

$$mr(a) = \frac{\sum_{i=1}^{n} (x_i - a)^i f_i}{\sum_{i=1}^{n} f_i}$$

مثال:

المشاهدات 5 ,4 ,3 ,2 ,1

احسب

a) $m_1(0)$ b) $m_3(4)$ c) $m_2(\overline{X})$

الحل:

a) $$M_1(0) = \frac{\sum_{i=1}^{5} (x_i)^1}{5}$$

$$= \frac{1+2+3+4+5}{5} = 3$$

b) $$M_3(4) = \frac{\sum_{i=1}^{5} (x_i - 4)^3}{5}$$

$$M_3(4) = \frac{(1-4)^3 + (2-4)^3 + (3-4)^3 + (4-4)^3 + (5-4)^3}{5}$$

$$M_3(4) = \frac{-27 - 8 - 1 + 0 + 1}{5}$$

$$= -7$$

c) $\bar{x} = 3$

$$M_2(\overline{X}) = \frac{\sum_{i=1}^{5} (x_i - 3)^2}{5}$$

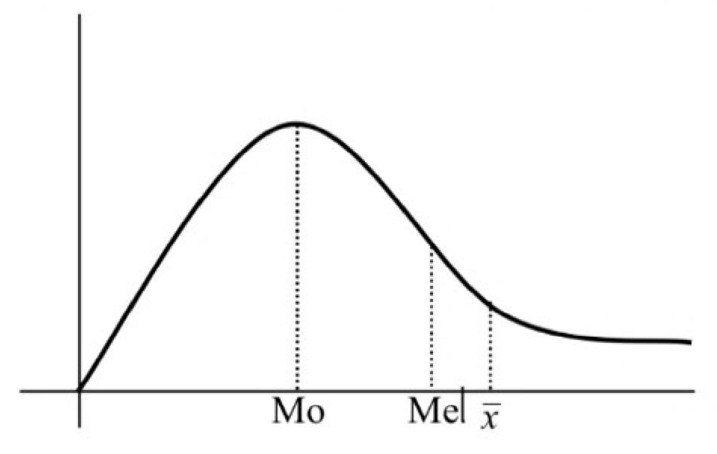
$$= \frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{5}$$

$$= \frac{4+1+0+1+4}{5}$$

$$= 2$$

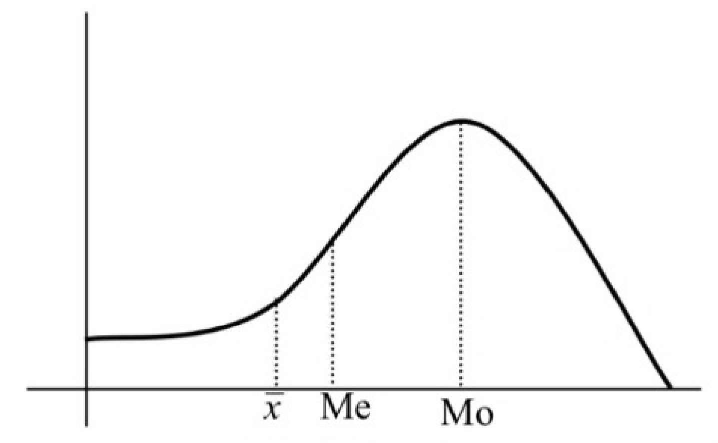
سنوضح الالتواء Skewness من خلال الأشكال الثلاثة التالية:

1- ملتوي نحو اليمين (موجب الالتواء) (Skewed to the right)



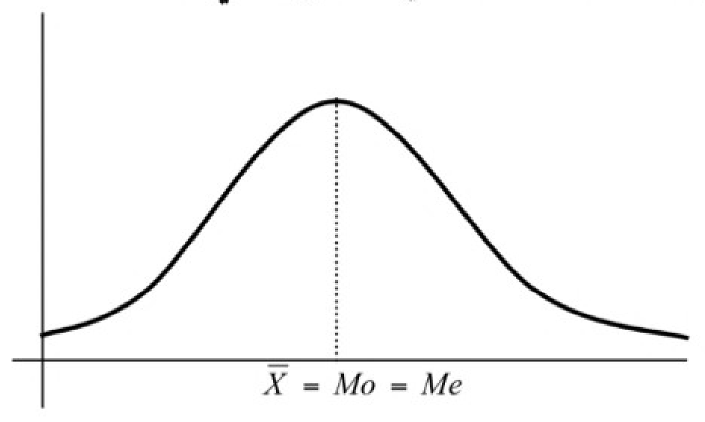
يكون المنوال ≥ الوسيط ≥ الوسط الحسابي

2- أما في التوزيعات الملتوية نحو اليسار (Skewed to the left) (سالبة الالتـواء) كمـا في الشكل التالي:



يكون الوسط الحسابي ≥ الوسيط ≥ المنوال

3- وفي حالة التوزيعات المتماثلة كما في الشكل التالي:



أي أن الوسط الحسابي = الوسيط = المنوال

تعريف: يعرف مقياس الالتواء بالقانون التالي:

$$L_3 = \frac{m_3(\bar{x})}{\left(\sqrt{m_2(\bar{x})}\right)^3}$$

ويوجد مقاييس أخرى للالتواء منها

$$\text{معامل الالتواء الربيعي} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

مثال:

للجدول التكراري التالي: احسب معامل الالتواء (L_3)

Class	10-14	15-19	20-24	25-29	30-34
Frequency	1	3	5	9	2

$$L_3 = \frac{m_3(\bar{x})}{\left(\sqrt{m_2(\bar{x})}\right)^3}$$

$$m_3(\bar{x}) = \frac{\sum (x_i - \bar{x})^3 f_i}{\sum f_i}$$

$$m_2(\bar{x}) = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}$$

Class make (x_i)	*Frequency* f_i	$f_i x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^3 f_i$
12	1	12	-12	-1728	-1728
17	3	51	-7	-343	-1029
22	5	110	-2	-8	-40
27	9	243	3	27	243
32	2	64	8	512	1024
Total	20	480			-1530

$$\overline{X} = \frac{480}{20} = 24$$

$$m_3(\overline{x}) = \frac{-1530}{20}$$

$$= -\ 76.5$$

$(x_i - \overline{x})^2$	$(x_i - \overline{x})^2 f_i$
144	144
49	147
4	20
9	81
64	128
	520

$$m_2(\overline{X}) = \frac{520}{20}$$

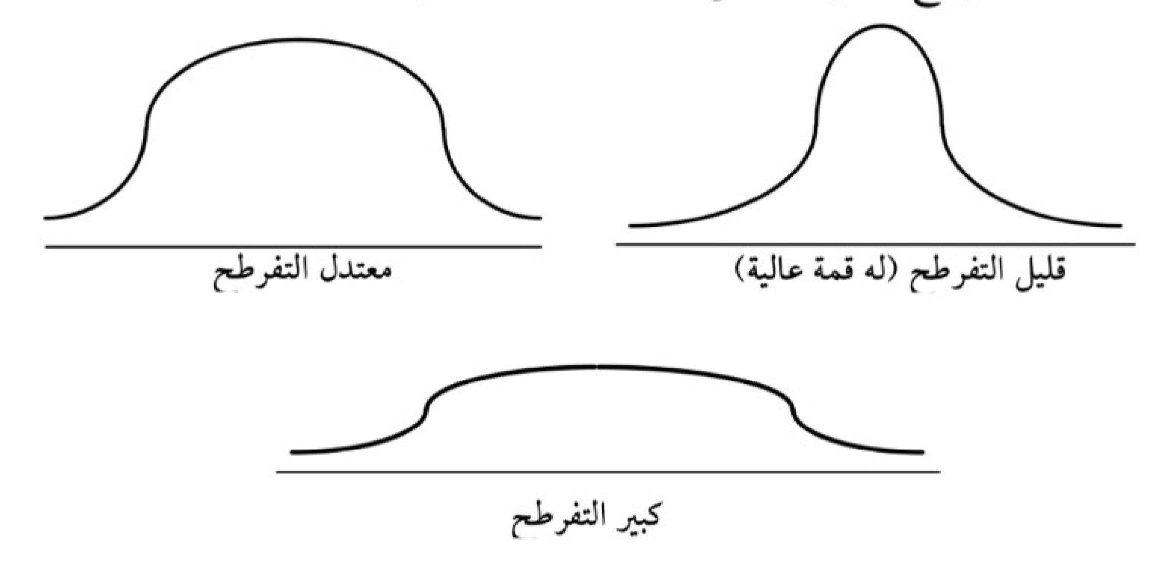
$$= 26$$

$$L_3 = \frac{-67.5}{\left(\sqrt{26}\right)^3}$$

$$= -0.58$$

تمرين: ماذا تعني الإشارة السالبة في الجواب؟

أما التفرطح فسنوضحه من خلال الأشكال التالية:



ومقياس التفرطح هو:

$$K_4 = \frac{m_4(\bar{x})}{(m_2(\bar{x}))^2}$$

تمرين: في المثال السابق جد قيمة (K_4)

ملاحظة عامة:

جميع مقاييس النزعة المركزية وهي الوسط الحسابي والوسيط والمـوال تتـأثر بجميع العمليات الحسابية الأربع فإذا ضربت كل مشاهدة بالعـدد الحقيقـي a ثـم أضيف العدد الحقيقي b فإن:

المقياس بعد التعديل = x a المقياس قبل التعديل +b

فإذا رمزنا للمقياس قبل التعديل (m_x) ورمزنا للمقياس بعد التعـديل (m_y) فإن

$$m_y = am_x + b$$

مثال:

إذا كان الوسيط للمفردات $x_1, x_2, \ldots, x_n$ يساوي (50) جـد الوسـيط للمفـردات $80-x_1, 80-x_2, \ldots, 80-x_n$

الحل:

$$Me = 80 - 50$$
$$= 30 \quad \text{(بعد التعديل)}$$

تمارين

1- إذا كان سعر السهم لشركة ما في سوق عمان المالي بالدينار على مدار أسبوع كما يلي: 5.5, 5.3, 5.2, 5.3, 5 , 5.1, 5.1

فأوجد:

أ- الوسط الحسابي لسعر السهم في ذلك الأسبوع $(\bar{x})$.

ب- الوسيط. (Me)

ج- المنوال. (Mo)

2- إذا اخترنا عينة عشوائية من الأسر حجمها (10) وأخذ دخل كل أسرة بالدينار فكانت 150, 170, 200, 250, 130, 120, 130, 100, 110, 190

فأوجد:

أ- الوسط الحسابي لدخل الأسرة في هذه العينة.

ب- الوسيط.

ج- المنوال.

3- الجدول التالي يبين فئات أطوال 50 طالباً في إحدى المدارس الأساسية:

Class	100-109	110-119	120-129	130-139	140-149	150-159	Total
Frequency	1	2	7	20	15	5	50

أوجد :

أ- الوسط الحسابي بطريقة الوسط الفرضي

ب- الوسيط.

ج- المنوال

د- المئين 60.

هـ- الرتبة المئينية للطول 125.

و- جد الطول الذي يزيد عنه 35 طالباً وماذا نسمي هذا الطول؟

4- إذا كان عدد الطلاب المتقدمين لمادة مبادئ الإحصاء (20) طالباً وعدد الطالبات (30) طالبة. فإذا علمت أن الوسط الحسابي للطلاب (70) والوسط الحسابي للطالبات (75). فاحسب الوسط الحسابي لجميع الطلبة المتقدمين لهذه المادة؟

5- إذا كانت أعمار (40) شخصاً موزعة كما في الجدول التالي:

Class	20-24	25-29	30-34	35-39	40-44	Total
Frequency	6	8	10	9	7	40

فأوجد

أ- الوسط الحسابي لأعمار هؤلاء الأشخاص.

ب-الوسيط.

ج- المنوال.

د- العشير 7.

هـ- الربيع الأدنى.

و- الربيع الأعلى.

ز- العشير الرابع بيانياً.

6- إذا كان الوسط الحسابي لخمسين طالبا هو (70). فراجع المعلم طالبان فزادت علامة الأول بمقدار (10) ونقصت علامة الآخر بمقدار (5). احسب الوسط الحسابي لجميع الطلبة بعد عملية المراجعة؟

7- إذا كان الوسط الحسابي لتوزيع أحادي المنوال هو 50 وكان المنوال هو (40). فأوجد الوسيط.

8- إذا كانت مبيعات أحد المتاجر لشهر نيسان من عام 1998. بمئات الدنانير هي:

4	7	5	9	12	17
6	16	11	18	7	19
12	10	14	20	8	13
4	20	13	14	11	6
9	8	19	18	13	15

أ- كوّن جدول توزيع تكراري بخمس فئات ومنه جد الوسط الحسابي والوسيط والمنوال.

ب- جد المئين 80 لمبيعات هذا المتجر.

9- إذا كانت الأوساط الحسابية لعلامات مادة الإحصاء لثلاث شعب أ، ب، جـ هي 70، 80، x على الترتيب وكانت أعداد هذه الشعب على التوالي 20، 30، 50، احسب قيمة (x) إذا علمت أن الوسط الحسابي المرجح لهذه الشعب هو 75.5؟

10- احسب مجموع علامات 50 طالباً الوسط الحسابي لعلاماتهم 85؟

11- إذا كانت انحرافات خمس قيم عن وسطها الحسابي هي a,8 – 2a, 2-a , 4a , 3a جد قيمة a؟

12- إذا كانت مجموع انحرافات خمسون قيمة عن الوسط الفرضي (20) هو (70). فجد الوسط الحسابي لهذه القيم؟

13- مستشفى فيه ثلاثون ممرضا وعشرة أطباء. إذا كان معدّل رواتب الممرضين (200) ديناراً ومعدّل رواتب الأطباء (700) ديناراً وقررت إدارة المستشفى زيادة رواتب الأطباء عشرون ديناراً. وزيادة رواتب الممرضين 10٪ من رواتبهم احسب:

أ- معدل رواتب الأطباء بعد الزيادة.

ب- معدّل رواتب الممرضين بعد الزيادة.

ج- معدّل رواتب الأطباء والممرضون معا قبل وبعد الزيادة.

14- إذا كانت $x_1, x_2, \ldots, x_n$ مجموعة من المشاهدات وسطها الحسابي $(\bar{X})$ فأثبـت أن

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0 .$$

15- إذا كان لدينا جدول تكراري فيـه m مـن الفئـات، مراكزهـا هـي $x_1, x_2, \ldots, x_m$ والتكـرارات المقابلـة لكـل فئـة هـي $f_1, f_2, \ldots, f_m$ علـى الترتيـب فأثبـت أن

$$\sum_{i=1}^{m}(x_i - \bar{x})f_i = 0$$

16- إذا كان لـدينا $x_1, x_2, \ldots, x_n$ مجموعـة مـن المـشاهدات، وسـطها الحـسابي $(\bar{X})$ ، وعدّلت هذه المشاهدات حسب العلاقة $y = ax+b$ حيث a, b عـددان حقيقيـان، x المشاهدة قبل التعديل، y المشاهدة بعد التعديل. فأثبت أن $\bar{y} = a\bar{x} + b$

17- إذا كان لدينا المشاهدات $x_1, x_2, \ldots, x_n$

بين أن:

$$m_1(0) = \bar{x}$$

18- للجدول الوارد في تمرين (5) جد ما يلي:

أ- $M_2(30)$

ب- $M_4(0)$

ج- L_3

3

الوحدة الثالثة

مقاييس التشتت

Measures of Dispersion

مقاييس التشتت

Measures of Dispersion

تستخدم مقاييس التشتت لإعطاء صورة عن مدى تقارب (تجانس) المشاهدات أو تباعدها (تشتتها) من بعضها البعض. فكلما زادت قيمة مقياس التشتت كلما ازداد تشتت المشاهدات وكلما قلت قيمته كلما زاد التجانس بين المشاهدات.

وجميع قيم مقاييس التشتت غير سالبة، ومن مقاييس التشتت المدى، نصف المدى الربيعي، الانحراف المتوسط، الانحراف المعياري.

أولاً: المـــدى Range

أ- للمشاهدات المفردة:

تعريف: المدى في حالة المفردات هو أكبر مشاهدة –أصغر مشاهدة والبعض يسميه المدى المطلق.

Range = max. Observation-min. observation

مثال:

أوجد المدى للمفردات 4, 5, 16, 10, 4, 5, 2,9.

الحل:

Range = 16 – 2 = 14

ملاحظة: المدى يتأثر بالقيم الشاذة "المتطرفة"، فبعض الأحيان لا يعطي صورة حقيقية عن واقع المشاهدات لهذا السبب.

ب- للجداول التكرارية:

تعريف:

Range = Upper boundary of the last class – lower boundary of the first class

مثال:

أوجد المدى للجدول التكراري التالي:

الفئات	19-10	29-20	39-30	49-40	59-50	المجموع
التكرارات	5	10	20	15	10	60

الحل:

Range = 59.5 – 9.5
= 50

ثانياً: نصف المدى الربيعي Semi-inter quartile range

$$\text{semi – inter quartile range} = \frac{Q_3 - Q_1}{2}$$

مثال:

أوجد نصف المدى الربيعي للجدول التالي:

Class	10-19	20-29	30-39	40-49	50-59	Total
Frequency	2	4	7	2	5	20

الحل: نحول الجدول التكراري إلى جدول تكراري تراكمي.

Upper boundary	Cumulative frequency
Less than 19.5	2
Less than 29.5	6
Less than 39.5	13
Less than 49.5	15
Less than 59.5	20

لإيجاد Q_3 :

$$C.R = \frac{75}{100} \times 20$$

$$=15$$

وبالنظر للجدول التكراري التراكمي نجد أن $Q_3 = 49.5$

لإيجاد Q_1:

$$C.R = \frac{25}{100} \times 20$$
$$= 5$$

بالنظر للجدول التكراري التراكمي نجد أن هذه القيمة تقع بين التكرارين التراكميين 2، 6

4	3	2 .... 19.5	10
		5 .... Q_1	
		6 .... 29.5	

$$Q_1 = 19.5 + \frac{3}{4} \times 10$$
$$= 19.5 + 7.5$$
$$= 27$$

$$\therefore \quad \text{semi – interquartile range} = \frac{49.5 - 27}{2}$$
$$= 11.25$$

ملاحظة: البعض يستخدم المدى الربيعي كمقياس تشتت بدلا من نصف المدى الربيعي.

حيث $\text{Quartile Range} = Q_3 - Q_1$

ثالثاً: الانحراف المتوسط Mean deviation

أ- في حالة المفردات

تعريف: إذا كانت $x_1, x_2, \ldots, x_n$، مجموعة من المفردات وسطها الحسابي $\bar{x}$ فيعرف الانحراف المتوسط (Mean Deviation M.D) بالعلاقة التالية:

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

وبشكل مختصر

$$M.D = \frac{\sum |x - \bar{x}|}{n}$$

مثال:

أوجد الانحراف المتوسط للمفردات 2, 4, 7, 5, 2.

الحل:

$$\bar{X} = \frac{2 + 4 + 7 + 5 + 2}{5}$$

$$\bar{X} = 4$$

$$\therefore M.D = \frac{|2 - 4| + |4 - 4| + |7 - 4| + |5 - 4| + |2 - 4|}{5}$$

$$= \frac{2 + 0 + 3 + 1 + 2}{5}$$

$$\text{M.D} = \frac{8}{5} = 1.6$$

ب- في حالة الجداول التكرارية

تعريف: إذا كانت $x_1, x_2, \ldots, x_n$ مراكـز فئـات جـدول تكـراري عـدد فئاتـه m. $f_1, f_2, \ldots, f_m$ التكرارات المقابلة لهذه الفئات على الترتيـب فيعـرف الانحـراف المتوسـط لهذا الجدول بالعلاقة التالية:

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}| f_i}{\sum_{i=1}^{n} f_i}$$

وبشكل مختصر

$$M.D = \frac{\sum |x - \bar{x}| f}{\sum f}$$

ملاحظة: نفس القانون الـسابق يـستخدم في حالـة المفـردات المتكـررة. بحيـث تمثـل x قيمة المفردة بدلاً من مركز الفئة.

مثال:

أوجد الانحراف المتوسط للجدول التكراري التالي:

class	5-9	10-14	15-19	20-24	25-29	Total
Frequency	3	2	5	6	4	20

الحل:

$$M.D = \frac{\sum |x - \bar{x}| f}{\sum f}$$

class	frequency	Class-mark (x)	xf	$\|x - \bar{x}\|$	$\|x - \bar{x}\| f$
5-9	3	7	21	11.5	34.5
10-14	2	12	24	6.5	13
15-19	5	17	85	1.5	7.5
20-24	6	22	132	3.5	21
25-29	4	27	108	8.5	34
Total	20		370		110

$$\bar{X} = \frac{370}{20}$$

$$= 18.5$$

$$M.D = \frac{110}{20}$$

$$= 5.5$$

مثال:

أوجد الانحراف المتوسط للجدول التالي:

Observation	5	6	8	15	20
frequency	6	3	4	4	3

الحل:

x	f	xf	$\lvert x-\bar{x}\rvert$	$\lvert x-\bar{x}\rvert f$
5	6	30	5	30
6	3	18	4	12
8	4	32	2	8
15	4	60	5	20
20	3	60	10	30
Total		200		100

$$\bar{X} = \frac{200}{20} = 10$$

$$M.D = \frac{100}{20} = 5$$

رابعاً: الانحراف المعياري (σ) Standard Deviation

أ- في حالة المفردات

تعريف: إذا كانت $x_1, x_2, \ldots, x_n$ مجموعة من المشاهدات وسطها الحسابي $\bar{x}$ فـإن الانحراف المعياري لها (يرمز له بالرمزσ) يعطى بالعلاقة

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$$

ويسمى مربع الانحراف المعياري σ^2 بالتباين. (Variance)

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n}$$

مثال:

أوجد الانحراف المعياري والتباين للمفردات 5 ,4 ,3 ,2 ,1

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$$

$$\bar{X} = \frac{1+2+3+4+5}{5}$$

$$= 3$$

$$\sigma = \sqrt{\frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{5}}$$

$$\sigma = \sqrt{\frac{4+1+0+1+4}{5}}$$

$$\therefore \sigma = \sqrt{2}$$

$$\sigma^2 = 2$$

مثال:

إذا كانت $x_1, x_2, \ldots, x_n$ مجموعة من المشاهدات فأثبت أن:

$$\sigma = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum x_i^2 - 2\bar{x}x_i + \bar{x}^2}{n}}$$

بتوزيع المجموع على الحدود نحصل على

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2}{n}}$$

لكن

$$\sum_{i=1}^{n} x_i = n\bar{x}$$

$$\therefore \sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - 2\bar{x}(n\bar{x}) + n\bar{x}^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n}}$$

وهو المطلوب.

ملاحظة: بينا أن

$$\sigma = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n}}$$

$$\Rightarrow \sigma = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$

$$\Rightarrow \sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

كذلك يمكننا كتابة الصيغة الأخيرة على الصورة

$$\sigma = \sqrt{\bar{x}^2 - (\bar{x})^2}$$

يمكننا تلخيص ما سبق بأن قانون الانحراف المعياري يمكن كتابته على عدة صور منها:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$\sigma = \sqrt{\bar{x}^2 - (\bar{x})^2}$$

ولكن الأكثر استخداماً الشكلين الأول والثالث.

تمرين:

أوجد الانحراف المعياري للمفردات 1, 2, 3, 4, 5 باستخدام الشكل الثاني لقانون الانحراف المعياري؟

مثال:

إذا كان لدينا سبع مشاهدات بحيث أن $\sum_{i=1}^{7} x_i^2 = 140$ ، $\sum_{i=1}^{7} x_i = 28$ فأوجد الانحراف المعياري لهذه المشاهدات؟

الحل:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{7} x_i^2}{7} - \left(\frac{\sum_{i=1}^{7} x_i}{7}\right)^2}$$

$$\sigma = \sqrt{\frac{140}{7} - \left(\frac{28}{7}\right)^2}$$

$$= \sqrt{20 - 16}$$

$$= 2$$

ب – في حالة الجداول التكرارية

ستتعرض لطريقتين في حساب الانحراف المعياري للجـداول التكراريـة إلا وهمـا الطريقـة العامـة وطريقـة الوسـط الفرضـي. وكـذلك تـستخدم الطـريقتين في حـساب الانحراف المعياري للمفردات المتكررة. باستخدام نفس القوانين.

1- الطريقة العامة لحساب الانحراف المعياري General method

تعريف: إذا كانت $x_1, x_2, \ldots, x_n$ مراكز فئات جدول تكـراري وكانـت تكـرارات فئاته $f_1, f_2, \ldots, f_m$ على الترتيب فإن

$$\sigma = \sqrt{\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2 f_i}{\sum_{i=1}^{m} f_i}}$$

الـصيغة الـواردة في التعريـف الـسابق للانحـراف المعيـاري تـؤدي إلى صـعوبة في العمليات الحسابية وخاصة إذا كان الوسط الحسابي عدد غير صحيح. لـذلك نـستخدم الصيغة التالية:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{m} x_i^2 f_i}{\sum_{i=1}^{m} f_i} - \left(\frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i}\right)^2}$$

تمرين:

أثبت أن الصيغتين السابقتين للانحراف المعياري في حالة الجداول التكرارية متكافئتين؟

يمكن كتابة الصيغة الثانية للانحراف المعياري في حالة الجداول التكرارية بالصورة

$$\sigma = \sqrt{\bar{x}^2 - (\bar{x})^2}$$

مثال:

أوجد التباين (Variance) للجدول التكراري التالي؟

Class	10-14	15-19	20-24	25-29	30-34	Total
Frequency	1	2	6	7	4	20

الحل:

$$\sigma^2 = \frac{\sum x^2 f}{\sum f} - \left(\frac{\sum xf}{\sum f}\right)^2$$

Class	frequency	Class-mark(*x*)	xf	X^2	X^2f
10-14	1	12	12	144	144
15-19	2	17	34	289	578
20-24	6	22	132	484	2904
25-29	7	27	189	729	5103
30-34	4	32	128	1024	4096
Total	20		495		12825

$$\sigma^2 = \frac{12825}{20} - \left(\frac{495}{20}\right)^2$$

$$= 641.25 - 612.5626$$

$$\text{Variance } \left(\sigma^2\right) = 28.6875$$

ويمكننا تبسيط الإجراءات الحسابية باستخدام طريقة الوسط الفرضي التالية:

2- طريقة الوسط الفرضي Assumed-mean method

إذا كان لدينا مجموعة من المشاهدات وأضفنا لها أو طرحنا منها مقداراً ثابتاً فإن ذلك لا يؤثر على تباعد القيم عن بعضها البعض. وذلك هو المبدأ الذي نستخدمه في حساب الانحراف المعياري بطريقة الوسط الفرضي.

إذا كانت $x_1, x_2, \ldots, x_m$ مراكز فئات جدول تكراري، بحيث أن التكرارات المقابلة للفئات هي $f_1, f_2, \ldots, f_m$ على الترتيب فإن قانون الانحراف المعياري بطريقة الوسط الفرضي هو

$$\sigma = \sqrt{\frac{\sum_{i=1}^{m} d_i^2 f_i}{\sum_{i=1}^{m} f_i} - \left(\frac{\sum_{i=1}^{m} d_i f_i}{\sum_{i=1}^{m} f_i}\right)^2}$$

حيث $d_i = x_i - A$

A: assumed-mean (الوسط الفرضي)

والوسط الفرضي تختاره مركز فئة متوسـطة الموقـع في الجـدول ويكـون تكرارهـا كبير نوعا ما.

وبشكل أكثر اختصاراً

$$\sigma = \sqrt{\frac{\sum df}{\sum f} - \left(\frac{\sum df}{\sum f}\right)^2}$$

مثال:

أوجد الانحراف المعياري للجدول التكراري التالي بطريقة الوسط الفرضي.

Class	10-19	20-29	30-39	40-49	50-59	Total
Frequency	1	2	3	3	1	10

الحل:

$$\sigma = \sqrt{\frac{\sum d^2 f}{\sum d} - \left(\frac{\sum df}{\sum d}\right)^2}$$

class	frequency	class mark (x)	d=x-34.5	df	d^2	d^2f
10-19	1	14.5	-20	-20	400	400
20-29	2	24.5	-10	-20	100	200
30-39	3	34.5	0	0	0	0
40-49	3	44.5	10	30	100	300
50-59	1	54.5	20	20	400	400
Total	10			10		1300

لاحظ أننا اعتبرنا A = 34.5

$$\therefore \sigma = \sqrt{\frac{1300}{10} - \left(\frac{10}{10}\right)^2}$$
$$= \sqrt{130 - 1}$$
$$= \sqrt{129}$$
$$= 11.36$$

معامل الاختلاف Coefficient of Variation

يعرف معامل الاختلاف بأنه النسبة المئوية بين الانحراف المعياري والوسط الحسابي ويرمز له بالرمز CV، حيث

$$C.V = \frac{\sigma}{\overline{X}} \times 100\,\%$$

مثال:

إذا كان الانحراف المعياري لتوزيع ما (5.33) والوسط الحسابي (24.75) احسب معامل الاختلاف لهذا التوزيع؟

الحل:

$$C.V = \frac{\sigma}{\overline{X}} \times 100\,\%$$
$$= \frac{5.35}{24.75} \times 100\%$$
$$= 21.61\,\%$$

ويستخدم معامل الاختلاف للمقارنة بين مجموعتين فكلما كان معامل الاختلاف أقل كان تجانس القيم أكثر.

مثال:

إذا كان الوسط الحسابي والانحراف المعياري لإنتاج مصنعين لمدة عشرة سنوات كما يلي:

المصنع A: $\overline{X} = 30{,}8 = 8$

المصنع B: $\overline{X} = 60{,}8 = 12$

أي المصنعين إنتاجه أفضل.

الحل:

$$C.V_A = \frac{8}{30} \times 100\% = 26.67\%$$

$$C.V_B = \frac{12}{60} \times 100\% = 20\%$$

بما أن معامل الاختلاف للمصنع الثاني أقل من المصنع الأول فإن إنتاج المصنع الثاني أفضل لأنه أكثر تجانساً.

تمرين:

أعد حل المثال السابق باستخدام الطريقة العامة؟

> **ملاحظة عامة:**
> جميع مقاييس التشتت لا تتأثر بإضافة ثابت حقيقي لجميع المفردات. ولكنها تتأثر بضرب جميع المفردات بثابت حقيقي. (انظر تمرين5).

تمارين

1- للمفردات التالية: 6 ,0 ,1- ,4 ,1

أوجد:

1. المدى range؟
2. الانحراف المتوسط mean-deviation؟
3. الانحراف المعياري standard deviation؟
4. التباين variance؟

2- أعد حل السؤال السابق للمفردات: 1, 7, 12, 2, 5, 17, 20, 15, 16, 18, 1, 1
4, 3, 7, 9

3- للجدول التكراري التالي:

Class	-10-(-1)	0-9	10-19	20-29	30-39	40-49	Totall
Frequency	5	7	9	8	10	1	40

أوجد:

1. المدى range ؟
2. نصف المدى الربيعي semi interquartile range؟
3. الانحراف المتوسط Mean deviation؟
4. التباين variance؟

4- في السؤال السابق أوجد المدى العشري للتوزيع حيث:

المدى العشيري = العشير التاسع – العشير الأول

Decile range = D_9-D_1

5- إذا كان لدينا المفردات $x_1, x_2,\ldots,x_n$، وعدّلت هذه المفردات حسب العلاقة $y=ax+b$، حيث a, b عددان حقيقيان. x المفردة قبل التعديل y المفردة بعد التعديل فأثبت أن:

أ- الانحراف المتوسط بعد التعديل = | a |× الانحراف المتوسط قبل التعديل.

ب- الانحراف المعياري بعد التعديل = | a |× الانحراف المعياري قبل التعديل.

6- إذا كان لدينا مجموعة من المفردات، انحرافها المعياري 5، وعدّلت جميع المفردات حسب العلاقة $y= 6-2x$ فأوجد:

أ- الانحراف المعياري بعد التعديل؟

ب- التباين قبل التعديل؟

ج- التباين بعد التعديل وما علاقته بالتباين قبل التعديل؟

7- إذا كان الانحراف المعياري لعشرة مشاهدات هو (5). أوجد مجموع مربعات انحرافات القيم عن وسطها الحسابي؟

8- إذا كان التباين للقيم 1, 5,k ,4- هو (11.5). أوجد الوسط الحسابي وقيمة k ؟ (كم حلا للسؤال؟).

9- إذا كان مجموع مربعات مئة قيمة هو (1500). ومجموع هذه القيم (300)، جد تباين هذه القيم؟

10- إذا كانت انحرافات ستة قيم عن وسطها الحسابي هي 0 ,6- ,7 ,2- ,5 ,4-. جد الانحراف المعياري لهذه القيم؟ وكذلك الانحراف المتوسط؟

11- إذا كان الانحراف المعياري والوسط الحسابي لرواتب مجموعة موظفين في شركة بالدينار (12)، (180) على الترتيب. فإذا قرر مدير الشركة زيادة الرواتب بمقدار عشرة دنانير لكل موظف. احسب الانحراف المعياري والوسط الحسابي للرواتب بعد التعديل؟

12- للجدول التكراري التالي

Class	100-119	120-139	140-159	160-179	180-199	Total
Frequency	17	13	24	15	11	80

احسب:

أ- الوسط الحسابي بطريقة الوسط الفرضي.

ب-الانحراف المعياري بطريقة الوسط الفرضي.

ج- المدى الربيعي.

د- المدى العشيري.

13- إذا كانت $x_1, \; x_2, \ldots, x_n$ مجموعة من القيم وكان $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 272$
والانحراف المعياري لهذه القيم (4) جد n؟

14- إذا كان نصف المدى الربيعي لمجموعة من القيم (7.5) والمئين 75 لها (23) جد الربيع الأول لهذه المشاهدات؟

15- إذا كان مجموع مربعات (65) قيمة هي (4875) وانحرافها المعياري $\sqrt{11}$.
احسب الوسط الحسابي لهذه القيم؟

16- للمفردات $x_1, x_2, \ldots, x_n$ بين أن:

أ- $\sigma^2 = m_2(\bar{x})$

ب- $L_3 = \dfrac{m_2(\bar{x})}{\sigma^2}$



الوحدة الرابعة

الارتباط والانحدار

Correlation and Regression

الارتباط والانحدار

Correlation and Regression

المقدمة

الارتباط: هو قوة العلاقة بين متغيرين أو أكثر. ولكن سيقتصر تعاملنا في هذا الكتاب على متغيرين فقط.

جداول الانتشار وعلاقتها بالارتباط Scatter Tables

الانتشار هو التمثيل البياني للعلاقة بين المتغيرين ويكون ذلك برصد نقاط المتغيرين على المحورين الأفقي والعمودي.

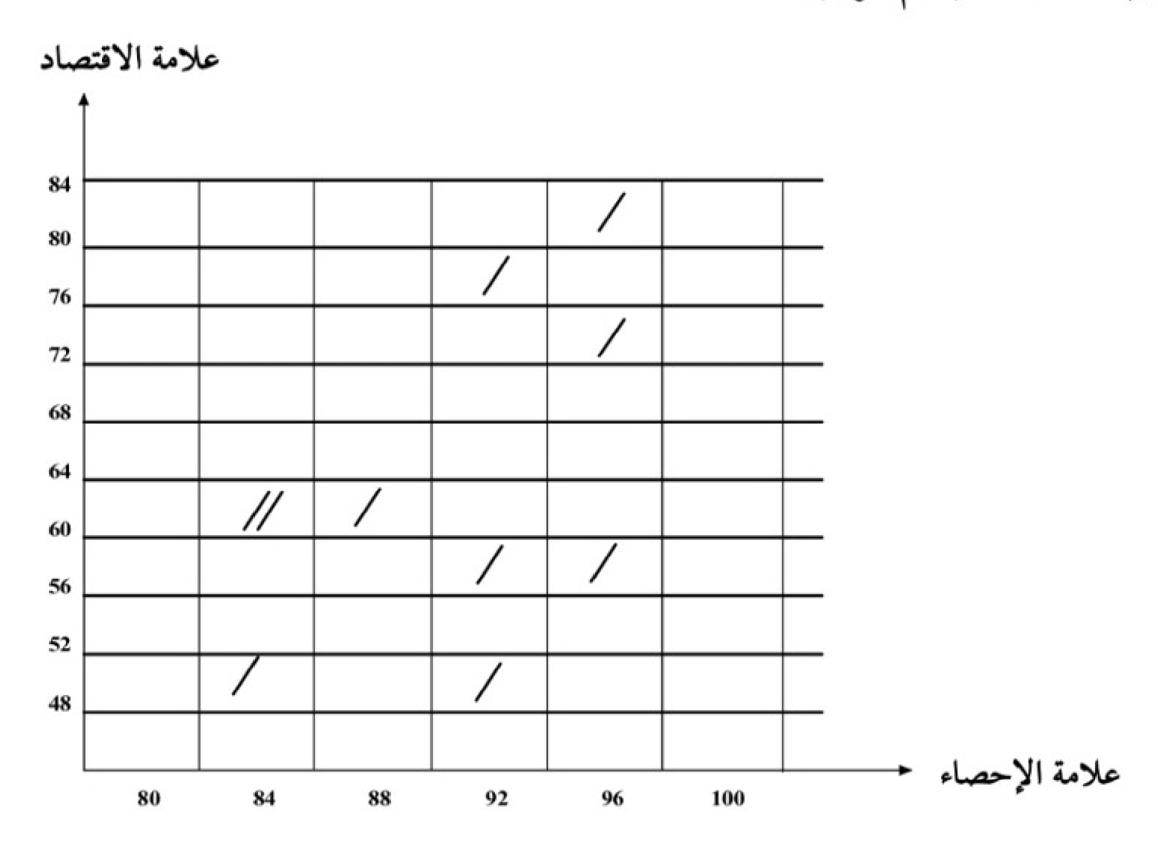
ومن خلال لوحة الانتشار يمكن تحديد مدى قوة العلاقة بين المتغيرين سواء كانت هذه العلاقة طردية (إيجابية) أو عكسية (سلبية).

مثال:

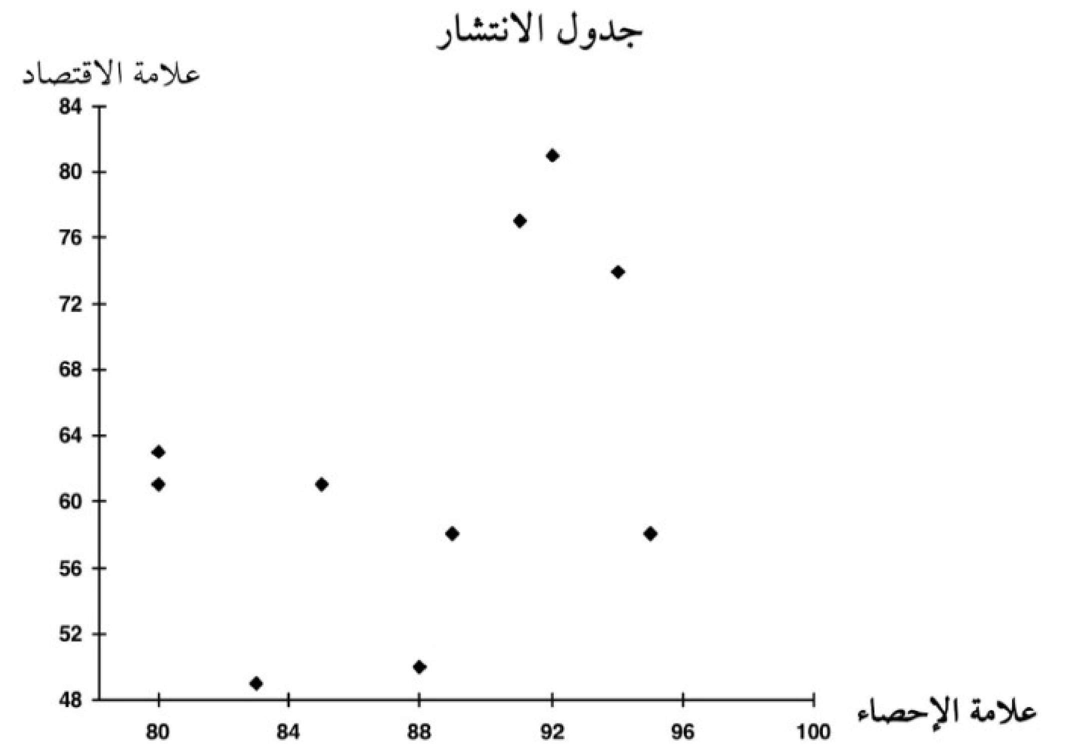
في دراسة أجريت لقياس مدى العلاقة بين التحصيل في مادة الإحصاء ومادة الاقتصاد، أجري امتحان في المادتين لعشرة طلاب وكانت النتائج كما يلي:

الإحصاء	الاقتصاد
80	61
95	58
94	74
89	58
92	81
83	49
91	77
88	50
80	63
85	61

كون جدول الانتشار ولوحة الانتشار وبين فيما إذا كانت العلاقة قوية أم ضـعيفة وإذا كانت سالبة أم موجبة.



جدول الانتشار



نرى من خلال لوحة الانتشار وجدول الانتشار أن الارتباط موجب ولكنه ضعيف نوعا ما ولكن إذا أردنا معرفة قوة العلاقة الحقيقية بين المتغيرين فإن أفضل مقياس هو معامل الارتباط والذي سنتحدث عنه في البند اللاحق.

معامل الارتباط Correlation Coefficient

وهو المقياس الرقمي لقوة الارتباط بين متغيرين.

خصائص معامل الارتباط Properties of correlation coefficient

1- تتراوح قيمة معامل الارتباط (r) بين 1- و 1.

أي $-1 \le r \le 1$

2- إذا كانت r بين 0 و (1) فإن العلاقة بين المتغيرين تكون علاقة موجبة أو طردية أما إذا كانت r بين 1- , 0 تكون العلاقة عكسية أو سالبة.

3- إذا كانت r =1 فإن العلاقة تكون موجبة تامة. وإذا كانت r = -1 فإن العلاقة سالبة تامة.

4- إذا كانت r = 0 فإنه لا يوجد علاقة بين المتغيرين.

5- تزداد قوة العلاقة كلما اقتربنا من (1) و (-1) وتقل كلما اقتربنا من 0.

وهناك عدة أنواع من معاملات الارتباط ولكن ستقتصر دراستنا في هذا الكتاب على معاملي ارتباط بيرسون وسبيرمان.

1- معامل ارتباط بيرسون Persons Correlation Coefficient: ويسمى معامل ارتباط العزوم. يعد من أفضل معاملات الارتباط وأكثرها شيوعاً.

$(x_1 , y_1) , (x_2, y_2), \dots, (x_n, y_n)$

فإن معامل ارتباط بيرسون هو:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n S_x S_y}$$

حيث:

$\bar{x}$ = الوسط الحسابي لمشاهدات المتغير x.

$\bar{y}$ = الوسط الحسابي لمشاهدات المتغير y.

σ_x = الانحراف المعياري للمتغير x.

σ_y = الانحراف المعياري للمتغير y.

مثال:

أوجد معامل ارتباط بيرسون بين قيم x,y من البيانات التالية وبين فيما إذا كانت العلاقة طردية أم عكسية، ضعيفة أم قوية.

X	150	162	180	160	170	180
Y	200	250	300	200	240	280

الحل:

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x}).(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
150	200	−17	−45	765	289	2025
162	250	−5	5	−25	25	25
180	300	13	55	715	169	3025
160	200	−7	−45	315	49	2025
170	240	3	−5	−15	9	25
180	280	13	35	455	169	1225
1002	1470			2210	710	8350

نجد أولاً : $\bar{x} = 167$, $\bar{y} = 245$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} = \sqrt{\frac{710}{6}} = 10.88$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}} = \sqrt{\frac{8350}{6}} = 37.31$$

$$r = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x \sigma_y} = \frac{2210}{6 * 10.88 * 37.31} = 0.907$$

وهذا يدل على أن العلاقة طردية قوية.

وهناك أشكال أخرى لمعادلة إيجاد معامل ارتباط بيرسون منها.

$$r = \frac{n\left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

$$r = \frac{\sum xy - n\overline{x}\overline{y}}{\sqrt{\sum x^2 - n\overline{x}^2}\sqrt{\sum y^2 - n\overline{y}^2}}$$

مثال:

احسب معامل ارتباط بيرسون بين علامات عشرة طلاب في مساقي الرياضيات والإحصاء من الجدول التالي:

x	80	60	55	40	75	85	70	60	80	90
y	75	65	60	50	70	90	70	55	80	85

الحل:

ستستخدم القانون

$$r = \frac{\sum xy - n\overline{x}\overline{y}}{\sqrt{\sum x^2 - n\overline{x}^2}\sqrt{\sum y^2 - n\overline{y}^2}}$$

الرياضيات (x)	الإحصاء (y)	xy	x^2	y^2
80	75	6000	6400	5625
60	65	3900	3600	4225
55	60	3300	3025	3600
40	50	2000	1600	2500
75	70	5250	5625	4900
85	90	7650	7225	8100
70	70	4900	4900	4900
60	55	3300	3600	3025
80	80	6400	6400	6400
90	85	7650	8100	7225
695	700	50350	50475	50500

$$\bar{x} = 69.5$$
$$\bar{y} = 70$$

$$r = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2}\sqrt{\sum y^2 - n\bar{y}^2}}$$

$$= \frac{50350 - 10 * 69.5 * 70}{\sqrt{50475 - 10(69.5)^2}\ \sqrt{50500 - 10(70)^2}}$$

$$= \frac{1700}{\sqrt{2172.5}\ \sqrt{1500}} = \frac{1700}{1805.2} = 0.942$$

الارتباط (طردي قوي)

2- معامل ارتباط سبيرمان Sparman Correlation Coefficient:

ويسمى معامل ارتباط الرتب

يستخدم إذا كان هناك صعوبة في استخدام معامل ارتباط بيرسون أو لم يتوفر لدينا القيم الحقيقية للمشاهدات ولكن توفرت رتبها فقط ويكون:

$$r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$

حيث di: فروق بين رتب المتغيرين x , y.

أي $d_i = O_x - O_y$

حيث رتبة x : Ox

رتبة y : Oy

وبشكل مختصر

$$r = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

ولإيجاد الرتب نجد رتب كل متغير على حده فمثلاً لإيجاد رتب المتغير x نعطي أكبر مشاهدة الرتبة (1) والتي تليها الرتبة (2) وهكذا. أما إذا تساوت أكثر من قيمة نعطي كل قيمة معدل رتب هذه القيم.

مثال:

أوجد معامل ارتباط الرتب للجدول التالي والذي يمثل رتب عشرة طلاب في موضعين دراسيين. (x, y).

Ox	6	7	10	8	9	1	2	5	3	4
Oy	8	5	9	7	10	2	6	4	1	3

الحل:

$$r = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

O_x	O_y	d_i	d_i^2
6	8	-2	4
7	5	2	4
10	9	1	1
8	7	1	1
9	10	-1	1
1	2	-1	1
2	6	-4	16
5	4	1	1
3	1	2	4
4	3	1	1
			34

$$r = 1 - \frac{6 * 34}{(10)((10)^2 - 1)} = 1 - \frac{204}{990} = 0.794$$ (طردية قوية)

مثال:

أوجد معامل ارتباط سبيرمان للجدول التالي والذي يمثل تكـاليف الدعايـة لنـوع من البضائع وقيمة المبيعات بمئات الدنانير:

تكاليف الدعاية (X)□	8	10	6	4	12	13	5	11	9
المبيعات (Y)□	150	160	150	130	165	180	120	160	150

الحل:

$$r = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

تكاليف الدعاية (X)	المبيعات (y)	O_x	O_y	d_i	di^2
8	150	6	6	0	0
10	160	4	3.5	0.5	0.25
6	150	7	6	1	1
4	130	9	8	1	1
12	165	2	2	0	0
13	180	1	1	0	0
5	120	8	9	−1	1
11	160	3	3.5	−0.5	0.25
9	150	5	6	−1	1
					4.5

$$r = 1 - \frac{6 * 4.5}{9(9^2 - 1)}$$

$= 1\text{-}0.0375$

$= 0.9625$ □

أثر التحويلات الخطية على معامل الارتباط:

إذا كان r معامل الارتباط بين المتغيرين x , y. وأجرينا تعديلاً على كل من قيم x, y كالآتي:

عددان حقيقيان a, b $X^* = ax + b\ ,\ a \neq 0$

عددان حقيقيان c, d $Y^* = cy + d\ ,\ c \neq 0$

فإن معامل الارتباط لا يتأثر إذا اتفقت a, c في الإشارة، وتتغير إشارته فقط إذا اختلفتا في الإشارة.

مثال:

إذا كان معامل الارتباط بين المتغيرين x,y هو r = 0.9 وعدلت قيم x وفق المعادلة $X^* = 2x+6$ كما عدلت قيم y وفق المعادلة $y^* = 8-3y$ فما هو معامل الارتباط بين x^* , y^* بعد التعديل.

الحل:

معامل الارتباط الجديد هو 0.9- وذلك لأن معامل x ومعامل y مختلفان في الإشارة.

مثال:

احسب معامل ارتباط بيرسون للجدول التالي:

X	38	41	36	34	37	36
Y	53	57	51	48	52	51

الحل:

نستطيع تعديل المشاهدات حسب المعادلتين $X^* = x-30$

$Y^* = y-45$ دون أن يؤثر ذلك على معامل الارتباط فتصبح القيم كالآتي:

$$r_{x,y} = r_{x^*,y^*}$$

$$r = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2}\ \sqrt{\sum y^2 - n\bar{y}^2}}$$

x^*	y^*	x^*y^*	$(x^*)^2$	$(y^*)^2$
8	8	64	64	64
11	12	132	121	144
6	6	36	36	36
4	3	12	16	9
7	7	49	49	49
6	6	36	36	36
42	42	329	322	338

$$\left(\bar{x}^*\right) = \frac{42}{6} = 7$$

$$\left(\bar{y}^*\right) = \frac{42}{6} = 7$$

$$r = \frac{329 - 6*7*7}{\sqrt{322 - 6*7^2}\ \sqrt{338 - 6*7^2}}$$

$$= \frac{35}{\sqrt{28}\ \sqrt{44}} = 0.997$$

تمريـن

أعد حل المثال باستخدام القيم الأصلية؟

الانحدار Regression

تعريف: معادلة خط الانحدار (regression line equation): هـي معادلـة خطيـة بين متغيرين x, y وتستخدم في التنبؤ بقيم متغير إذا عرف المتغير الآخر.

وهناك صورتان لمعادلة خط الانحدار وهما:

أ- معادلة خط انحدار Y على X وهي:

$$Y = b + ax$$

حيث $a = \frac{\sigma_y}{\sigma_x} r$ ويسمى (a) معامل الانحدار

σ_x = الانحراف المعياري لقيم x.

σ_y = الانحراف المعياري لقيم y.

r = معامل الارتباط بين x, y

ويمكن الحصول على قيمة a باستخدام طريقة المربعات الصغرى لتكون

$$a = \frac{\sum xy - n\overline{x}\overline{y}}{\sum x^2 - n\overline{x}^2}$$

أما قيمة b فنجدها من المعادلة $\overline{y} = a\overline{x} + b$ أي أن

$$b = \overline{y} - a\overline{x}$$

والمعادلة y = b+ax تستخدم للتنبؤ بقيمة y إذا علمت قيمة x.

مثال:

حسب معامل الارتباط بين نتائج الطلبة في الامتحان (x) والامتحان (y) فكانت r = 0.8 ، وكانت $\overline{x} = 59$ ، $\sigma_x=14$ ، $\overline{y} = 66$ ، $\sigma_y=7$ فأوجد معادلة خط انحدار y على x ثم أوجد نتيجة الطالب المتوقعة في الامتحان y إذا حصل على علامة 65 في الامتحان x.

الحل:

$$y = b + ax$$

نحسب قيمة a من المعادلة

$$a = \frac{\sigma_y}{\sigma_x} r$$

$$= \frac{7}{14}.(0.8) = 0.4$$

$$b = 66 - 0.4\ (59)$$

$$= 42.4$$

فتكون معادلة خط الانحدار هي:

$$y = 42.4 + 0.4x$$

نتيجة الطالب المتوقعة في الامتحان y إذا كانت علامته في الامتحان 65 = x تساوي (65) هي

$$\hat{y} = 42.4 + 0.4 \times 65 = 68.4$$

مثال:

في شركة لتجارة السيارات يمثل الجدول التالي عدد السيارات المباعة x في السنوات 2000-2004 والربح y بآلاف الدنانير.

X	50	40	45	55	60
Y	75	63	50	72	80

فأوجد:

1- معادلة خط انحدار y على x.

2- لو ا فترضنا أن الشركة ستبيع 50 سيارة في عام 2005 فما الربح المتوقع لها في هذه السنة.

الحل:

1- لإيجاد معادلة خط الانحدار $y = b+ax$

نجد أولاً قيمة a, b

x	y	xy	x^2
50	75	3750	2500
40	63	2520	1600
45	50	2250	2025
55	72	3960	3025
60	80	4800	3600
250	340	17280	12750

$$\bar{x} = \frac{250}{5} = 50$$

$$\bar{y} = \frac{340}{5} = 68$$

$$a = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

$$= \frac{17280 - 5 \times 50 \times 68}{12750 - 5(50)^2}$$

$$= \frac{280}{250} = 1.12$$

$$b = \bar{y} - a\bar{x}$$

$$= 68 - 1.12 \times 50$$

$$= 68 - 56 = 12$$

∴ معادلة خط الانحدار

$$y = 12+1.12x$$

2- الربح المتوقع للشركة إذا باعت 50 سيارة في عام 2005 هو:

$$\hat{y} = 12 + 1.12x$$

$$= 12 + 1.12 \times 50 = 68$$

أي 68000 J.D

ب – معادلة خط انحدار X على Y وهي:

$$x = d + cy$$

$$c = \frac{\sigma_x}{\sigma_y} r$$

$$c = \frac{\sum xy - n\bar{x}\bar{y}}{\sum y^2 - n\bar{y}^2}$$

$$d = \bar{x} - c\bar{y}$$

وتستخدم معادلة خط انحدار x على y للتنبؤ بقيمة x إذا علمت قيمة y.

مثال:

إذا كانت معادلة خط الانحدار للمتغيرين y،x هي x = 106+5y وكانـت $\sigma_x=18$، $\sigma_y=3$ ، فما هو معامل الارتباط بين x,y.

الحل:

نعلم من معادلة خط انحدار x على y أن

$$c = \frac{\sigma_x}{\sigma_y}.r$$

$$\Rightarrow r = c\frac{\sigma_y}{\sigma_x}$$

$$= 5 * \frac{3}{18} = \frac{15}{18} = 0.833$$

مثال:

في امتحان تحصيلي لستة طلاب في مادتي الرياضـيات (x) والإحـصاء (y) كانـت النتائج كالتالي:

الرياضيات (x)	60	75	50	40	63	72
الإحصاء (y)	80	83	55	70	60	78

أوجد :

أ- معادلة خط انحدار x على y.

ب- إذا حصل طالب على علامة (60) في الإحصاء فماذا تكون علامته في الرياضيات.

الحل:

أ- معادلة خط انحدار x على y هي = x= d+cy

حيث

$$c = \frac{\sum xy - n\overline{x}\overline{y}}{\sum y^2 - n\overline{y}^2}$$

الرياضيات (x)	الإحصاء (y)	xy	y^2
60	80	4800	6400
75	83	6225	6889
50	55	2750	3025
40	70	2800	4900
63	60	3780	3600
72	78	5616	6084
360	426	25971	30898

$$\bar{x} = \frac{360}{6} = 60$$

$$\bar{y} = \frac{426}{6} = 71$$

$$c = \frac{\sum xy - n\bar{x}\bar{y}}{\sum y^2 - n\bar{y}^2}$$

$$= \frac{25971 - 6 * 60 * 71}{30898 - 6(71)^2}$$

$$= \frac{411}{652} = 0.63$$

$$d = \bar{x} - c\bar{y}$$

$$= 60 - 0.63 \times 71$$

$$= 15.27$$

$\therefore$ معادلة خط الانحدار هي $x = 15.27 + 0.63y$

ب- إذا حصل الطالب على علامة (60) في الإحصاء (y) فإن علامته المتوقعة في الرياضيات (x) تكون:

$$\hat{x} = 15.7 + 0.63 \times 60$$

$$= 53.07$$

تعريف: الخطأ في التنبؤ = القيمة الحقيقية – القيمة المتنبأ بها

e= actual vale (θ) - estimated value of θ

i.e: $e = \theta - \hat{\theta}$

ففي المثال السابق الفرع (ب)

يكون الخطأ في التنبؤ

$$e = x - \hat{x}$$

$$e = 63-53.07$$

$$= 9.93$$

مثال:

إذا كانت معادلة خط انحدار y علىx هي $y = 10+0.7x$ وكانت قيمة y التي تقابل القيمة $(x=20)$ هي 25 فما هو الخطأ في التنبؤ بقيمة y.

الحل:

قيمة y المتنبأ بها هي: $\hat{y} = 10+0.7 \times 20$

$=24$

الخطأ في التنبؤ

$$e = y - \hat{y}$$

ملاحظات:

إذا كانت (x_1, y_1) , (x_2, y_2) … (x_n, y_n) مجموعة من الأزواج المرتبة لقيم x، y.
واستخدمت لإيجاد:

أ- معادلة انحدار y على x وهي $y = b+ax$

ب- معادلة انحدار x على y وهي $x = d+cy$

فإنه

1- إذا مثلت المعادلتين في نفس المستوى البياني تكون نقطة تقاطع خطي الانحدار هي $(\bar{x}, \bar{y})$.

2- a, c لهما نفس الإشارة وهي إشارة معامل الارتباط r.

3- $r^2 = a * c \Rightarrow r = \pm\sqrt{ac}$

مثال:

إذا كانت معادلة انحدار y على x هي y =8-0.9x وكانت معادلة انحـدار x علـى y هي x= 9-0.4y، احسب معامل الارتباط بين المتغيرين x، y؟

الحل:

$$r^2 = -0.9 \times -0.4$$
$$= 0.36$$
$$\Rightarrow r = \pm 0.6$$
$$\Rightarrow r = -0.6$$

وقد أخذت الإشارة السالبة لأن معاملي الانحدار سالبين.

معامل التحديد The Coefficient of Determination

يمثل معامل التحديد نـسبة لانخفـاض في الأخطـاء عنـد اسـتخدام معادلـة خـط الانحدار، وتفسر تباين المشاهدات التي تفسر بمعادلـة خـط الانحـدار ويرمـز لـه بـالرمز (R^2) ونجده باستخدام القانون:

$$R^2 = \frac{a\left(\sum \chi_i y_i - n\bar{x}\bar{y}\right)}{\sum y_i^2 - n\bar{y}^2}$$

وتكون قيمته محصورة بين صفر و (1).

مثال:

إذا كانت معادلة خط الانحدار هي:

$$y = 12 + 1.12x$$

وكان $\sum \chi_i y_i = 17280, \bar{x} = 50, \bar{y} = 68$

$$\sum y^2 = 23678\ , n = 5$$

جد معامل التحديد

الحل:

$$R^2 = \frac{1.12(17280 - 5(50)(86))}{23678 - 5(68)^2}$$

$$R^2 = \frac{313.6}{558} = 0.562$$

تمرين: اثبت أن معامل التحديد (R^2) = مربع ارتباط بيرسون

تمارين

1- يمثل الجدول التالي الأطوال (x) والأوزان (y)لعشرة طلاب في إحدى كليات المجتمع.

X cm	170	172	165	175	168	180	160	158	173	167
Y kg	70	72	73	72	70	77	68	65	71	70

أ- أرسم لوحة الانتشار.

ب-احسب معامل ارتباط بيرسون بين أطوال وأوزان الطلبة.

ج- احسب معامل ارتباط سبيرمان بين الأطوال والأوزان.

د- أوجد معادلة خط انحدار y على x.

2- إذا كان لدينا المجاميع التالية:

$$\sum_{i=1}^{30} x_i = 1134 \qquad \sum_{i=1}^{30} y_i = 1154$$

$$\sum_{i=1}^{30} x_i^2 = 44564 \qquad \sum_{i=1}^{30} y_i^2 = 45636$$

$$\sum_{i=1}^{30} x_i y_i = 44931$$

a- احسب معامل ارتباط بيرسون.

b- أوجد معادلة خط انحدار x على y.

3- إذا كان معامل الارتباط بين x، y هو r = 0.7 وعدلت حسب المعادلتين

$x^* = 2.5x+7$

$y^* = 3y-6$

فما هو معامل الارتباط بين x^*, y^*.

4- إذا كان معامل ارتباط سبيرمان (الرتب) بين متغيرين x، y هو r = 0.8 وكان عدد أزواج المشاهدات يساوي 40، فأوجد مجموع مربعات فروق الرتب بين x,y.

5- إذا كانت معادلة خط الانحدار $y= 7+0.5x$ وكان

$$\sum_{i=1}^{10}(y_i - \bar{y})^2 = 250 \qquad \sum_{i=1}^{10}(x_i - \bar{x})^2 = 640$$

أوجد معامل الارتباط (r) بين x,y.

6- للجدول التالي:

X	2	4	6	8	10
Y	15	9	12	6	3

a- جد معامل ارتباط بيرسون بين قيم x، y؟

b- معادلة انحدار y على x؟

c- معادلة انحدار y على x؟

d- ارسـم معـادلتي الانحـدار في نفـس المـستوى وتأكـد أن نقطـة تقاطعهمـا هـي $(\bar{x}, \bar{y})$ ؟

e- هـ- احسب الخطأ في التنبؤ بقيمة y إذا علمت أن x=8؟

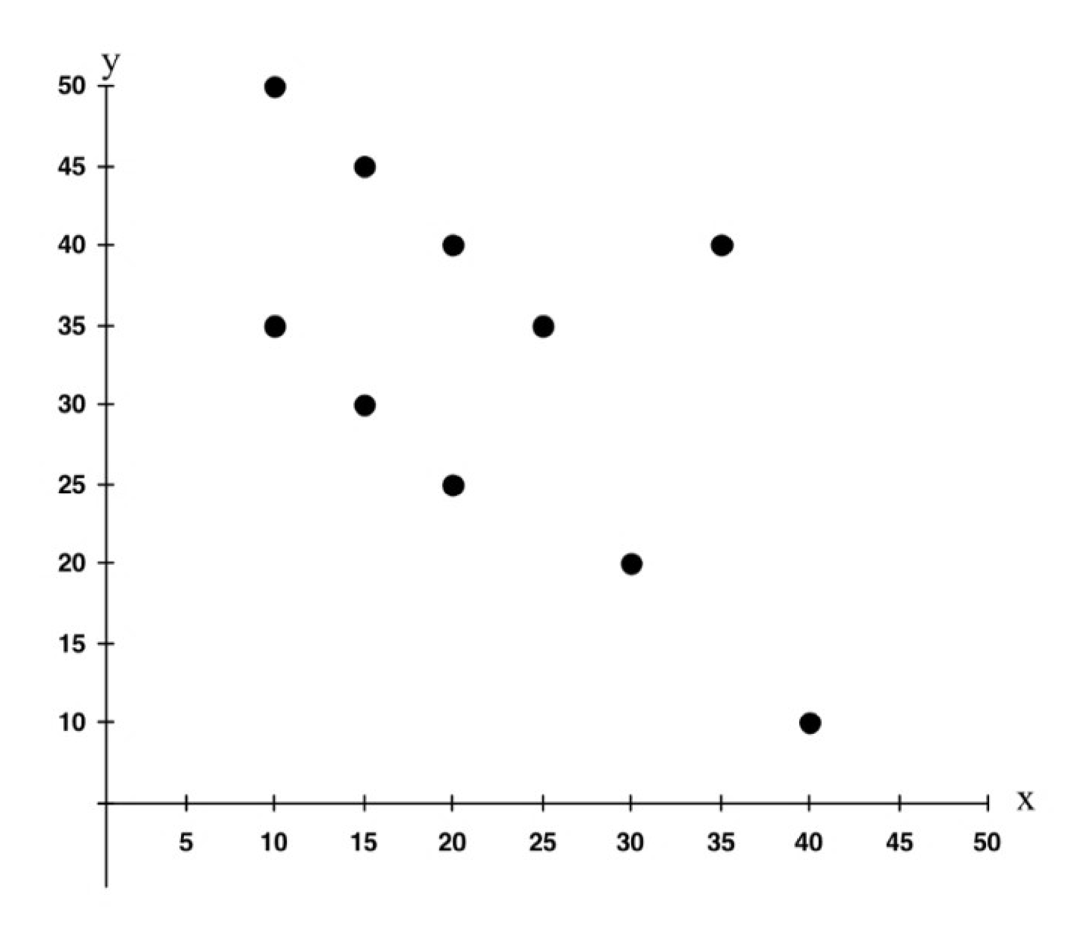
f- احسب الخطأ في التنبؤ بقيمة x إذا علمت أن قيمة y=6؟

7- إذا كانت معادلة انحدار y علـى x هـي $y = -15+2x$ وكانـت $r = 0.8$، $\bar{x} = 50$، فجد معادلة انحدار x على y؟

8- إذا كانت معادلة انحدار y على x هي $y =-2+1.2x$ ومعادلة انحدار x على y هـي $x = 11+0.7y$ جد الوسط الحسابي لكل من قيم x, y؟

9- إذا حُسب معامل ارتباط سبيرمان باستخدام الأزواج المرتبة (x_1, y_1) , (x_2, y_2) … (x_n, y_n)فكان (0.4). وكان مجموع مربعات فروق الرتـب لقيم x، y هو (99)، فما قيمة n؟

10- الشكل التالي هو شكل الانتشار لقيم x,y.



جد:

أ – معامل ارتباط بيرسون.

ب– معامل ارتباط سبيرمان.



الوحدة الخامسة

نظرية الاحتمال

Probability Theory

نظرية الاحتمال

Probability Theory

نظرية الاحتمالات تهتم بما يسمى بالتجارب العشوائية، والتجارب العشوائية هي تلك التجارب التي يمكن حصر نتائجها مسبقاً ولكن لا يمكن الجزم ماذا ستكون النتيجة. فمثلاً إذا ألقينا قطعة نقد فنكون عالمين سلفاً بأن النتيجة قد تكون صورة أو كتابة لكننا لا نجزم على أنها صورة أو أنها كتابة. والآن نبدأ بعرض بعض مفاهيم نظرية الاحتمالات.

الفضاء العيني (أو الفراغ العيني) والحادث Sample space and Event

تعريف: الفضاء العيني هو مجموعة كافة النتائج المتوقعة للتجربة العشوائية. وسنرمز للفضاء العيني بالرمز Ω "ويقرأ أوميغا".

الآن سنعطي أمثلة على فراغات عينية شائعة في موضوع الاحتمالات، لتكون مرجعية للقارئ عند حاجته لها.

1- تجربة رمي قطعة نقد مرة واحدة تكون $\Omega = \{H , T\}$

حيث

صورة H: Head

كتابة T: Tail

2- تجربة رمي قطعتي نقد مرة واحدة (أو قطعة مرتين)

$$\Omega= \{(H , H) , (H , T) , (T , H) , (T , T)\}$$

3- تجربة رمي ثلاث قطع نقدية مرة واحدة (أو قطعة ثلاث مرات).

$$\Omega = \{(H,H,H) , (H,H,T) , (H,T,H) , (T,H,H), (H,T,T), (T,H,T), (T,T,H), (T,T,T)\}$$

4- رمي حجر نرد مرة واحدة $\Omega = \{1,2,3,4,2,5,6\}$

5- رمي حجر نرد وقطعة نقد.

$\Omega = \{(1,H), (2,H), (3,H), (4,H), (5,H), (6,H), (1,T), (2,T), (3,T), (4,T), (5,T), (6,T)\}$

6- رمي حجري نرد مرة واحدة (أو رمي حجر مرتين)

$$\Omega = \{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6) \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6) \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6) \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6) \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$$

تعريف:

1- إذا كانت Ω فضاء عينيا لتجربة عشوائية ما فإن أية مجموعة جزئية مثل E من Ω($E \subset \Omega$) تدعى حادثا(event).

2- إذا كان E حادثا في Ω يتكون من عنصر واحد فإنه يسمى حادثا بسيطاً Simple event.

3- وإذا كان E حادثا فيه أكثر من عنصر من عناصر Ω فيسمى حادثا مركباً Compound event.

4- $\phi \subset \Omega$ وبالتالي فإن ϕ حادثا في Ω، يسمى هذا الحادث بالحادث المستحيل. (Null event)

5- $\Omega \subset \Omega$ وبالتالي فإن Ω حادثا في Ω، يسمى هذا الحادث بالحادث الأكيد.(Sure event).

بالرجوع إلى الفراغ العيني Ω لتجربة رمي قطعتي نقد إذا كان

$E_1 = \{(H,T), (T,H)\}$ حادث في Ω فإنه حادث مركب

أما $E_2 = \{(T,T)\}$ حادث بسيط في Ω.

أيضاً يمكنك إعطاء أمثلة على حوادث مختلفة بسيطة ومركبة لنفس التجربة.

مثال:

في تجربة رمي حجري النرد ليكن

E_1: هو الحادث الذي يكون مجموع الوجهين الظاهرين فيه (4).

E_2: هو الحادث الذي يكون مجموع الوجهين الظاهرين فيه (10).

E_3: هو الحادث الذي يكون الوجه الأول فيه فرديا (odd) والوجه الثاني زوجي (even).

E_4: هو الحادث الذي يكون مجموع الوجهين الظاهرين فيه أكبر من (10).

E_5: الوجه الأول يقبل القسمة على (divisible by)الوجه الثاني.

اكتب هذه الحوادث صريحة.

الحل:

$E_1= \{(1,3) , (3,1), (2,2)\}$

$E_2= \{(4,6) , (6,4), (5,5)\}$

$E_3= \{(1,2), (1,4) , (1,6) , (3,2), (3,4) , (3,6) , (5,2) , (5,4) , (5,6)\}$

$E_4= \{(5,6) , (6,5), (6.6)\}$

$E_5= \{(1,1), (2,1), (3,1) , (4,1), (5,1), (6,1), (2,2), (3,3), (4,4) , (5,5), (6,6), (4,2), (6,2), (6,3)\}$

تعريف: إذا كان E_1، E_2حادثين في الفراغ العيني Ω بحيث لا يوجد بينهما عناصر مشتركة أي $E_1 \cap E_2=\phi$ فيسمى الحادثين بحادثين منفصلين (disjoint events).

مثال:

في تجربة رمي حجر النرد مرة واحدة إذا كان

$E_1 = \{1,2,3\}$ $E_2=\{1,3,5\}$ $E_3=\{2,4\}$

فأي أزواج الحوادث التالية منفصلة:

a) E_1,E_2 b) E_1,E_3 c) E_2,E_3

الحل:

a- $E_1\cap E_2= \{1,3\}$ وبالتالي E_1,E_2 ليسا منفصلين.

b- $E_1\cap E_3= \{2\}$ وبالتالي E_1,E_3 ليسا منفصلين.

c- $E_2\cap E_3 = \phi$ وبالتالي E_2,E_3 منفصلان.

التكرار النسبي والاحتمال (Relative Frequency and Probability)

تعريف:

إذا كان n عدد مرات إجراء تجربة عشوائية، m عدد مرات الحصول على الحـادث E فإن $\frac{m}{n}$ تدعى بالتكرار النسبي للحادث أو الاحتمال التجريبي للحادث.

وعندما تصبح قيمة n كبيرة جداً ($n\to\infty$) فإن التكرار النـسبي يقـترب مـن قيمـة محددة سنرمز لها بالرمز p(E) وتسمى الاحتمال النظري للحادث E.

نشاط:

لتكن n عدد مرات إلقاء قطعة نقد، m عدد مرات ظهور الصورة.

أكمل الفراغ في الجدول التالي بعد إجراء التجربة عمليا.

عدد مرات إجراء التجربة (n)	عدد مرات الحصول على الصورة (m)	التكرار النسبي $\frac{m}{n}$
10	....	....
20	....	....
50	....	....
100	....	....
200	....	....

في النشاط السابق لو أصبحت قيمة (n) كبيرة جداً فإن $\left(\frac{m}{n}\right)$

ستقترب من القيمة $\frac{1}{2}$، وبالتالي فإن $P(\{H\}) = \frac{1}{2}$

تعريف:

إذا كانت Ω فضاء عينيا وكان كل حادث بسيط في Ω له نفس فرصة الحدوث. وكان E حادثا في Ω فإن

احتمال الحادث $(E) = \frac{\text{عدد عناصر } E}{\text{عدد عناصر } \Omega}$ ويسمى هذا الاحتمال النظري للحادث E:

$$P(E)=\frac{\#(E)}{\#(\Omega)}$$

حيث # تعني عدد العناصر.

وتسمى Ω في هذه الحالة بالفضاء العيني المنتظم (Uniform space).

تعريف: لأي فضاء عيني Ω (منتظم أو غير منتظم) يعرّف اقتران الاحتمال

$P:P(\Omega)\to \mathbf{R}$، حيث $P(\Omega)$ مجموعة كافة حوادث Ω، **R** مجموعة الأعداد الحقيقية.

من خلال المسلمات التالية:

1- لأي حادث $E\subset\Omega$ $P(\Omega)=1\ ,\quad P(E)\geq 0,$

2- إذا كان E_1، E_2 حادثين منفصلين في Ω فإن

$$P(E_1\cup E_2)= P(E_1)+P(E_2)$$

3- إذا كان E_1، E_2، ... حوادث في Ω بحيث

$$E_i\cap E_j=\phi$$

لأي $i\neq j$

$$P(E_1\cup E_2,\ldots)=P(E_1)+P(E_2)+\ldots$$

ملاحظة: تقرأ $P(E)$ احتمال الحادث E. (Probability of E)

خواص الاحتمال (Properties of probability):

إذا كانت Ω فضاء عينا فإنه

1- إذا كان E_1، E_2 حادثين في Ω، $E_1\subset E_2$ فإن $P(E_1)\leq P(E_2)$

2- لأي حادث E، $0\leq P(E)\leq 1$

3- $P(\phi)=0$

4- $P(E_1-E_2)=P(E_1)-P(E_1\cap E_2)$

5- $P(E_1\cap \overline{E}_2)=P(E_1)-P(E_1\cap E_2)$

مثال:

في تجربة رمي حجر النرد مرة واحدة إذا كان:

$E_1 = \{1,2\}$

$E_2 = \{1,4,5\}$

$E_3 =$ أن يكون الوجه الظاهر عدداً فردياً فأوجد

(a) $P(E_1)$ (b) $P(E_2)$ (c) $P(E_3)$
(d) $P(E_1 \cup E_2)$ (e) $P(E_1 \cap E_2)$

الحل:

لاحظ أن عدد عناصر $\Omega = 6$

a- $P(E_1) = \frac{2}{6} = \frac{1}{3}$

b- $P(E_2) = \frac{3}{6} = \frac{1}{2}$

c- $P(E_3) = \frac{3}{6} = \frac{1}{2}$

d- $E_1 \cup E_2 = \{1,2,4,5\} \Rightarrow P(E_1 \cup E_2) = \frac{4}{6} = \frac{2}{3}$

e- $E_1 \cap E_2 = \{1\} \Rightarrow P(E_1 \cap E_2) = \frac{1}{6}$

مثال:

صندوق يحوي (7) كرات سوداء، 3 كرات بيضاء، سحبت من الصندوق كرة واحدة عشوائياً (Raudomly)أوجد:

a- احتمال أن تكون سوداء (Black).

b- احتمال أن تكون بيضاء (White).

c- احتمال أن تكون حمراء (Red).

الحل:

a- $P(Black) = \frac{7}{10}$

b- $P(White) = \frac{3}{10}$

c- $P(Red) = p(\phi) = 0$

مثال:

إذا كان $P(E_1)=0.4$ ، $P(E_1 \cap E_2) = 0.3$ جد $P(E_1 - E_2)$

الحل:

$$P(E_1 - E_2) = P(E_1) - P(E_1 \cap E_2)$$
$$= 0.4 - 0.3$$
$$= 0.1$$

تعريف: إذا كان E حادثا في الفراغ العيني Ω، فعدم وقوع E يعرف بأنه متمم الحادث E (Complement of E) ويرمز له بالرمز $\overline{E}$ ويكون:

$$\mathbf{P(\overline{E}) = 1 - P(E)}$$

مثال:

إذا كان $P(E)=0.8$

فأوجد $P(\overline{E})$

الحل:

$$P(\overline{E}) = 1 - P(E)$$
$$= 1 - 0.8$$
$$= 0.2$$

تعريف: إذا كانت $E_1, E_2, \ldots, E_n$ حوادث في Ω بحيث أن

1- $E_1 \cup E_2 \cup \ldots \cup E_n = \Omega$

2- $E_i \cap E_j = \phi$ لكل $i \neq j$

فإن هذه الحوادث تسمى حوادث متباعدة وشاملة.

مثال:

في تجربة رمي حجر النرد مرة واحدة، إذا كان

$E_1 = \{1,6\}$ ، $E_2 = \{3,4,5\}$ ، $E_3 = \{3\}$

فهل هذه الحوادث متباعدة وشاملة؟

الحل:

$E_1 \cap E_2 = \phi$ ، $E_1 \cap E_3 = \phi$ ، $E_2 \cap E_3 = \phi$

أي أن E_1، E_2، E_3 منفصلة مثنى. (Mutualy exclusive)

$\Omega = E_1 \cup E_2 \cup E_3$

$\therefore$ E_1، E_2، E_3 حوادث متباعدة وشاملة.

نظرية: إذا كانت $E_1, E_2, \ldots, E_n$ حوادث متباعدة وشاملة فإن:

$$\mathbf{P(E_1) + P(E_2) + \ldots + P(E_n) = 1}$$

مثال:

إذا كانت E_1, E_2, E_3, E_4 حوادث متباعدة وشاملة وكان

$P(E_1) = 0.2,\ P(E_2) = 0.5,\quad P(E_3) = 0.1$

فأوجد $P(E_4)$؟

الحل:

E_1, E_2, E_3, E_4 حوادث متباعدة وشاملة

$$\therefore P(E_1) + P(E_2) + P(E_3) + P(E_4) = 1$$
$$0.2 + 0.5 + 0.1 + P(E_4) = 1$$
$$P(E_4) = 0.2$$

قانون جمع الاحتمالات

نظرية: إذا كان E_1, E_2 حادثين في Ω، فإن احتمال وقوع (E_1) أو (E_2) هو

$$\mathbf{P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)}$$

حيث $P(E_1 \cap E_2)$ هي احتمال حدوث الحادثين معا.

مثال:

إذا كان $P(A)=0.4$ ، $P(B)=0.8.$ ، $P(A \cap B) = 0.35$

فأوجد $P(A \cup B)$

الحل:

$$P(A\cup B)= P(A) + P(B)-P(A\cap B)$$
$$=0.4 + 0.8 - 0.35$$
$$= 0.85$$

مثال:

إذا كانت نسبة الأشخاص الذين عيونهم سوداء في مدينة عمان 70٪ وكانت نسبة الأشخاص الذين شعرهم أسود 40٪، ونسبة الأشخاص الذين عيونهم سوداء وشعرهم أسود 30٪. فإذا اخترنا شخصا عشوائيا من مدينة عمان فأوجد احتمال:

a- أن تكون عيونه سوداء أو شعره أسود.

b- أن لا يكون ذو عيون سوداء.

c- أن يكون ذو شعر أسود وعيونه ليست سوداء.

الحل:

لنفرض أن الحادث E_1 هو الحادث الذي يكون فيه الشخص الذي تم اختياره ذو عيون سوداء.

E_2 هو الحادث الذي يكون فيه الشخص الذي تم اختياره ذو شعر أسود.

فيكون $P(E_1)= 0.7$, $P(E_2)=0.4$, $P(E_1\cap E_2)=0.3$

a- $P(E_1\cup E_2) = P(E_1) + P(E_2) - P(E_1\cap E_2)$
$=0.7 + 0.4 - 0.3 = 0.8$

b- $P(\overline{E}_1) = 1- P(E_1)$
$= 1 - 0.7$
$= 0.3$

c- $P(\overline{E}_1 \cap E_2) = P(E_2) - P(E_1 \cap E_2)$
$= 0.4 - 0.3$
$= 0.1$

مثال:

إذا كان E_1,E_2 حادثين منفصلين في Ω وكان

$P(E_1)=\frac{3}{7}\ , P(E_2)=\frac{2}{7}$

فأوجد $P(E_1 \cup E_2)$

الحل:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$
$$= P(E_1) + P(E_2)$$
$$P(E_1 \cup E_2) = \frac{3}{7} + \frac{2}{7} = \frac{5}{7}$$

الحوادث المستقلة (Independent Events) "قانون ضرب الاحتمالات"

تعريف: إذا كان E_1 ، E_2 حادثين في Ω فيدعى E_1, E_2 حادثين مستقلين إذا كان أحدهما لا يتأثر بوقوع الآخر. وبصفة رياضية يكون E_1, E_2 مستقلين إذا كان

$$\mathbf{P(E_1 \cap E_2) = P(E_1)P(E_2)}$$

مثال:

إذا كـــان E_1 ، E_2 حـــادثين في Ω، وكـــان $P(E_1)=0.9$ ، $P(E_2)=0.5$ ، $P(E_1 \cap E_2)=0.45$ فهل E_1 ، E_2 مستقلان؟

الحل:

$$P(E_1)P(E_2) = (0.9)(0.5)$$
$$= 0.45$$
$$= P(E_1 \cap E_2)$$

إذن E_1 ، E_2 حادثين مستقلين.

مثال:

إذا كان E_1 ، E_2 حادثين مستقلين في Ω وكان $P(E_1)=0.6$ ، $P(E_2)=0.5$

فأوجد:

a- $P(E_1 \cap E_2)$

b- $P(E_1 \cup E_2)$

الحل:

$$a-\ P(E_1 \cap E_2) = P(E_1)P(E_2)$$
$$= (0.6)\ (0.5)$$
$$= 0.3$$

$$b-\ P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$
$$= 0.6+0.5 - 0.3$$
$$= 0.8$$

مثال:

أطلق صيادان نحو هدف فإذا كان احتمال إصابة الأول للهدف 0.8 واحتمال إصابة الثاني للهدف 0.6 فأوجد احتمال

a- إصابة الاثنين معا للهدف.

b- إصابة الهدف.

الحل:

E_1 : أن يصيب الأول الهدف $\Leftarrow$ $P(E_1) = 0.8$

E_2 : أن يصيب الثاني الهدف $\Leftarrow$ $P(E_2) = 0.6$

$$a-\quad P(E_1 \cap E_2) = P(E_1)P(E_2)$$
$$= (0.8)\ (0.6)$$
$$= 0.48$$

$$b-\quad P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$
$$= 0.8 + 0.6 - 0.48$$
$$= 0.92$$

مثال:

صندوق يحوي (6) كرات سوداء (Black)، (4) كرات حمراء (Red) سحب من الصندوق كرتان على التوالي مع الإرجاع، أوجد احتمال.

a- أن تكون الكرتان سوداوين.

b- أن تكون الكرة الأولى سوداء والثانية حمراء.

c- أن تكون إحدى الكرتين سوداء.

d- أن تكون الكرتان من نفس اللون.

الحل:

السحب هنا على التوالي مع الإرجاع فذلك يعني أن السحبة الثانية لا تتأثر بالسحبة الأولى فهذا يعني حوادث مستقلة.

(B)	(R)
6	4

a- $P(B, B) = \frac{6}{10} \times \frac{6}{10}$

$= \frac{36}{100}$

b- $P(B, R) = \frac{6}{10} \times \frac{4}{10}$

$= \frac{24}{100}$

c- P(إحدى الكرتين سوداء) $= P(B, R) + P(R, B)$

$= \frac{6}{10} \times \frac{4}{10} + \frac{4}{10} \times \frac{6}{10}$

$= \frac{48}{100}$

d- P (الكرتان نفس اللون) $= P(B, B) + P(R, R)$

$= \frac{6}{10} \times \frac{6}{10} + \frac{4}{10} \times \frac{4}{10}$

$= \frac{52}{100}$

تمرين: أعد حل المثال السابق إذا كان السحب دون إرجاع.

نظرية: إذا كان E_1 ، E_2 حادثين مستقلين في Ω فإن

1- E_1 ، $\overline{E}_2$ مستقلان.

2- $\overline{E}_1$ ، E_2 مستقلان.

3- $\overline{E}_1$ ، $\overline{E}_2$ مستقلان.

مثال:

إذا كان E_1 ، E_2 حادثين مستقلين في Ω ، $P(E_1) = 0.3$ ، $P(E_2) = 0.4$

فأوجد:

a- $P(E_1 \cap E_2)$

b- $P(\overline{E}_1 \cap \overline{E}_2)$

c- $P(\overline{E}_1 \cap E_2)$

d- $P(\overline{E}_1 \cup E_2)$

الحل:

a- $P(E_1 \cap E_2) = P(E_1)P(E_2)$
$= (0.3)\ (0.4)$
$= 0.12$

b- $P(\overline{E}_1 \cap \overline{E}_2) = P(\overline{E}_1)P(\overline{E}_2)$
$= (0.7)\ (0.6)$
$= 0.42$

c- $P(\overline{E}_1 \cap E_2) = P(\overline{E}_1)\,P(E_2)$
$= (0.7)\ (0.4)$
$= 0.28$

d- $P(\overline{E}_1 \cup E_2) = P(\overline{E}_1) + P(E_2) - P(\overline{E}_1 \cap E_2)$
$= 0.7 + 0.4 - 0.28$
$= 0.82$

الاحتمال المشروط ونظرية بيز

(Conditional Probability and Bay's Theorm)

تعريف:

إذا كان E_1 ، E_2 حادثين في Ω فـإن احتمـال حـدوث E_1 بـشرط حـدوث E_2 يرمز له بالرمز

$$P(E_1/E_2)=\frac{P(E_1\cap E_2)}{P(E_2)} \quad , \quad P(E_2)\neq 0$$

مثال:

إذا كان $P(E_1) = 0.7$ ، $P(E_2) = 0.8$ ، $P(E_1 \cap E_2) = 0.6$

فأوجد:

a) $P(E_1/E_2)$ b) $P(E_2/E_1)$ c- $P(E_1/\overline{E}_2)$

d) $P(\overline{E}_1/E_2)$ e) $P(\overline{E}_1/\overline{E}_2)$

الحل:

a- $$P(E_1/E_2)=\frac{P(E_1\cap E_2)}{P(E_2)}$$

$$=\frac{0.6}{0.8}$$

$$=\frac{3}{4}$$

b- $$P(E_2/E_1)=\frac{P(E_1\cap E_2)}{P(E_1)}$$

$$=\frac{0.6}{0.7}$$

$$=\frac{6}{7}$$

c- $$P(E_1/\overline{E}_2)=\frac{P(E_1\cap\overline{E}_2)}{P(\overline{E}_1)}$$

$$=\frac{P(E_1)-P(E_1\cap E_2)}{1-P(E_2)}$$

$$=\frac{0.7-0.6}{1-0.8}$$

$$=\frac{1}{2}$$

d- $$P(\overline{E}_1/E_2)=\frac{P(\overline{E}_1\cap E_2)}{P(E_2)}$$

$$=\frac{P(E_2)-P(E_1\cap E_2)}{P(E_2)}$$

$$P(\overline{E}_1/E_2)=\frac{0.8-0.6}{0.8}$$

$$=\frac{1}{4}$$

e- $$P(\overline{E}_1/\overline{E}_2)=\frac{P(\overline{E}_1\cap\overline{E}_2)}{P(\overline{E}_2)}$$

$$=\frac{P(\overline{E_1\cup E_2})}{P(\overline{E}_2)}$$ ☐ قانون ديمورغان

$$\therefore\quad P(\overline{E}_1/\overline{E}_2)=\frac{1-P(E_1\cup E_2)}{1-P(E_2)}$$

$$=\frac{1-0.9}{1-0.8}=\frac{1}{2}$$

مثال:

إذا كان $P(E_2)=0.45$ ، $P(E_1/E_2)=0.3$

أوجد

a- $P(E_1 \cap E_2)$

b- $P(\overline{E}_1 \cup \overline{E}_2)$

الحل:

a- $P(E_1 \cap E_2) = P(E_1/E_2)P(E_2)$

$= 0.3 \times 0.45$

$= 0.135$

b- $P(\overline{E}_1 \cup \overline{E}_2) = P(\overline{E_1 \cap E_2})$

$= 1 - P(E_1 \cap E_2)$

$= 1 - 0.135$

$= 0.865$

مثال:

إذا كان احتمال قبول صفاء في جامعة البلقاء 0.9 واحتمال قبول هيفاء في نفس الجامعة 0.8 واحتمال قبول الاثنتين معا 0.75 احسب:

a- احتمال قبول صفاء إذا قبلت هيفاء.

b- إذا قبلت صفاء فما احتمال قبول هيفاء.

c- إذا لم تقبل صفاء فما احتمال قبول هيفاء.

الحل:

E_1: أن تقبل صفاء $\Leftarrow$ $P(E_1) = 0.9$

E_2: أن تقبل هيفاء $\Leftarrow$ $P(E_2) = 0.8$

$\therefore P(E_1 \cap E_2) = 0.75$

□

a- $$P(E_1/E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$$
$$= \frac{0.75}{0.8}$$
$$= \frac{15}{16}$$

b- $$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$
$$= \frac{0.75}{0.9}$$
$$= \frac{5}{6}$$

c- $$P(E_2/\overline{E}_1) = \frac{P(\overline{E}_1 \cap E_2)}{P(\overline{E}_1)}$$
$$= \frac{P(E_2) - P(E_1 \cap E_2)}{1 - P(E_1)}$$
$$= \frac{0.8 - 0.75}{1 - 0.9}$$
$$= 0.5$$

ملاحظة: إذا كان E_1 ، E_2 حادثين مستقلين في Ω فإن

$$P(E_1/E_2) = P(E_1)$$
$$P(E_2/E_1) = P(E_2)$$

مثال:

صندوق يحوي ست كرات بيضاء وأربع كرات سوداء سحب من الصندوق كرتين على التوالي دون إرجاع احسب:

a- احتمال أن تكون الثانية سوداء إذا كانت الأولى بيضاء.

b- احتمال أن تكون الأولى بيضاء والثانية سوداء.

c- احتمال أن تكون الكرتان مختلفتان في اللون.

الحل:

سوداء	بيضاء
(B)	(W)
4	6

a- $P(B/W)=\frac{4}{9}$

b- $P(W\cap B)=P(B/W)P(W)$

$$=\frac{4}{9}\times\frac{6}{10}$$

$$=\frac{4}{15}$$

c- p (مختلفتان في اللون)$=P(B,W)+P(W,B)$

$$=\frac{4}{10}\times\frac{6}{9}+\frac{6}{10}\times\frac{4}{9}$$

$$=\frac{8}{15}$$

نظرية بيز Bay's Theorem

إذا كانت E_1 ، E_2 ، ...، E_n حوادث متباعدة وشاملة في الفضاء العيني Ω.

$E\subset\Omega$ فإن:

1. $P(E)=P(E/E_1)P(E_1)+P(E/E_2)P(E_2)+\ldots+P(E/E_n)P(E_n)$

2. $p(E_m/E)=\frac{P(E/E_m)P(E_m)}{P(E)}$, $m=1,2,\ldots,n.$

مثال:

يذهب رجل إلى عمله مستخدماً إحدى الوسائل التالية باص، سيارة، قطار مستخدما هذه الوسائل بنسبة مئوية 60٪، 30٪، 10٪ من الأيام على التوالي، فإذا كان احتمال أن يتأخر عن عمله إذا استخدم الباص 15٪ من الأيام وإذا استخدم

السيارة 8٪، وإذا استخدم القطار 20٪. فإذا اخترنا أحد الأيام التي يذهب فيها إلى العمل عشوائياً:

a- احسب احتمال أن يتأخر عن عمله في ذلك اليوم؟

b- إذا كان متأخراً عن عمله في ذلك اليوم ما احتمال أن يكون قد استخدم القطار؟

الحل:

E_1 : أن يستخدم الباص $\Leftarrow$ $P(E_1) = 0.6$

E_2 : أن يستخدم السيارة $\Leftarrow$ $P(E_2) = 0.3$

E_3 : أن يستخدم القطار $\Leftarrow$ $P(E_3) = 0.1$

E: أن يتأخر الرجل عن عمله.

$$P(E/E_1) = 0.15 \quad , \quad P(E/E_2) = 0.08 \quad , \quad P(E/E_3) = 0.2$$

a- $$P(E) = P(E/E_1)P(E_1) + P(E/E_2)P(E_2) + P(E/E_2)P(E_3)$$

$$= (0.15)\,(0.6) + (0.08)\,(0.3) + (0.2)\,(0.1)\square$$

$$\therefore \ P(E) = 0.134$$

b- $$P(E_3/E) = \frac{p(E/E_3)P(E_3)}{P(E)}$$

$$P(E_3/E) = \frac{(0.2)(0.1)}{0.134}$$

$$= \frac{10}{67}$$

مثال:

صندوقانA، B متشابهان، في A خمس كرات بيضاء وثلاث كرات سوداء، في B تسع كرات بيضاء وأربع كرات سوداء. فإذا اختير أحد الصندوقان عشوائيا ثم سحب منه كرة عشوائيا احسب احتمال أن تكون من الصندوق A إذا كانت بيضاء.

الحل:

E_1 : اختيار الصندوق A $\Leftarrow P(E_1) = \frac{1}{2}$

E_2 : اختيار الصندوق B $\Leftarrow P(E_2) = \frac{1}{2}$

E : أن تكون الكرة بيضاء.

$$P(E/E_1) = \frac{5}{8}$$

$$P(E/E_2) = \frac{9}{13}$$

$$P(E_2/E) = \frac{p(E/E_2)P(E_2)}{P(E)}$$

$$= \frac{p(E/E_2)P(E_2)}{P(E/E_1)P(E_1) + P(E/E_2)P(E_2)}$$

$$= \frac{\frac{9}{13} \times \frac{1}{2}}{\frac{5}{8} \times \frac{1}{2} + \frac{9}{13} \times \frac{1}{2}}$$

$$= \frac{72}{137}$$

المتغيرات العشوائية (Random Variables)

تعريف: المتغير العشوائي هو اقتران (Function) (X) مجاله الفضاء العيني (Sample Space) Ω ومداه X(Ω) هو مجموعة جزئية من الأعداد الحقيقية.

أي أن $X: \Omega \rightarrow \mathbf{R}$

وإذا كان مداه مجموعة جزئية من الأعداد النسبية **Q** يدعى متغيراً عشوائياً منفصلاً (discrete random variable).

أما إذا كان مداه يحوي فترة من الأعداد الحقيقية **(R)** فيدعى متغيراً عشوائياً متصلاً (continuous random variable).

والآن سندرس كل نوع من المتغيرات العشوائية على حده.

المتغير العشوائي المنفصل (Discrete Random Variable)

تعريف: إذا كان X متغيراً عشوائياً منفصلاً فإن الاقتران f(x) يدعى اقتران احتمال (Probability Function)أو اقتران توزيع (Distribution Function) إذا كان

$f(x) = P\ (X= x) \qquad x \in X\ (\Omega)$

وتسمى المجموعة $\{(x,f(x)): x \in X\ (\Omega)\}$ بالتوزيع الاحتمالي (Probability distribution).

مثال:

في تجربة رمى ثلاث قطع نقد، إذا عرّفنا المتغير العشوائي X على أنه عدد الصور (Heads) الظاهرة، أوجد

a- X (H,H,T)

b- X (T,T,T)

c- (مدى X) X(Ω)

الحل:

a- X (H,H,T) = 2
b- X (T,T,T) = 0
c- X(Ω) = {0,1,2, 3}

مثال:

كيس فيه ثلاث كرات بيضاء (White Balls) وأربع كـرات سـوداء (Black Balls). فـإذا سـحب مـن الكـيس خمـس كـرات علـى التـوالي (one after another) دون إرجـاع (without replacement). فإذا عرفنـا المـتغير العـشوائي X علـى أنـه عـدد الكـرات السوداء، أوجد مدى X

الحل:

$$X(\Omega) = \{2,3,4\}$$

مثال:

في تجربة رمي قطعتي النقد، إذا عرفنا المتغير العشوائي X علـى أنـه عـدد الـصور الظاهرة، اكتب التوزيع الاحتمالي للمتغير العشوائي.

الحل:

التوزيع الاحتمالي = $\{(x,f(x)) : x\in X(\Omega)\}$

$$X(\Omega) = \{0,1,2\}$$

فيكون التوزيع الاحتمالي

$$\left\{\left(0,\frac{1}{4}\right)\ ,\ \left(1,\frac{1}{2}\right)\ ,\ \left(2,\frac{1}{4}\right)\right\}$$

ملاحظة: إذا كان $f(x)$ اقتران احتمـال (Probability Function) للمـتغير العـشوائي X الذي مداه $X(\Omega)$ فيكون

1. $f(x) \geq 0 \qquad x\in X(\Omega)$
2. $\sum_{x\in X(\Omega)} f(x) = 1$

تعريف:

1- إذا كان X متغيراً عشوائياً فإن الوسط الحسابي (Mean) للمتغير العشوائي (X) هو

$$E(X) = \sum_{x\in X(\Omega)} X\ f(x)$$

كذلك يسمى الوسط الحسابي بالتوقع (Expectation) للمتغير العشوائي X.

2- يعرف العزم K (k^{th} moment) بأنه

$$E\left(X^{k}\right)=\sum_{x\in X(\Omega)} X^{k} f(x)$$

3- الانحراف المعياري (Standard deviation) للمتغير العشوائي (X) هو

$$\sigma=\sqrt{E(X^{2})-(E(X))^{2}}$$

ويسمى مربع الانحراف المعياري

$$\sigma^{2}=E(X^{2})-(E(X))^{2}$$

بالتباين (Variance).

مثال:

إذا كان X متغير عشوائياً يأخذ قيمة في المجموعة $x(\Omega)=\{1,2,3\}$ وكان اقتران الاحتمال له هو

$$f(x)=\frac{x}{6}$$

احسب:

1- التوزيع الاحتمالي.
2- توقع X
3- العزم الثاني للمتغير العشوائي X
4- Variance (x)

الحل:

1) التوزيع الاحتمالي: $\left\{\left(1,\frac{1}{6}\right),\left(2,\frac{1}{3}\right),\left(3,\frac{1}{2}\right)\right\}$

2) $$E\left(X\right)=\sum_{x=1}^{3} Xf(x)$$

$$E(X) = 1f(1) + 2f(2) + 3f(3)$$

$$= \frac{1}{6} + \frac{4}{6} + \frac{9}{6} = \frac{7}{3}$$

3) $E(X^2)$ = العزم الثاني

$$E\left(X^2\right) = \sum_{x=1}^{3} X^2 f(x)$$

$$= (1)^2 f(1) + (2)^2 f(2) + (3)^2 f(3)$$

$$= \frac{1}{6} + \frac{8}{6} + \frac{27}{6}$$

$$= \frac{36}{6}$$

$$= 6$$

4) $\sigma^2 = E(X^2) - (E(X))^2$ □

$$= 6 - \left(\frac{(7)}{3}\right)^2$$

$$= 6 - \frac{49}{9}$$

$$= \frac{54 - 49}{9}$$

$$= \frac{5}{9}$$

مثال:

في تجربة رمي حجري النـرد (Tossing two die) إذا عرّفنـا المـتغير العـشوائي X على أنه مجموع الوجهين الظاهرين،أوجد التالي:

1- X (Ω)

2- E(X)

3- Variance (X)

الحل:

1) $X(\Omega) = \{2,3,4,5,6,7,8,9,10,11,12\}$

2) $$E(X) = \sum_{x=2}^{12} Xf(x)$$

$$= (2)\left(\frac{1}{36}\right) + (3)\left(\frac{6}{36}\right) + (4)\left(\frac{3}{36}\right) + (5)\left(\frac{4}{36}\right)$$

$$+ (6)\left(\frac{5}{36}\right) + (7)\left(\frac{6}{36}\right) + (8)\left(\frac{5}{36}\right) + (9)\left(\frac{4}{36}\right)$$

$$+ (10)\left(\frac{3}{36}\right) + (11)\left(\frac{2}{36}\right) + (12)\left(\frac{1}{36}\right)$$

$$= 7$$

3) Exercise

تمرين:

في تجربة رمي حجري النرد، إذا عرّفنا المتغير العشوائي X على أنه الفرق المطلق للوجهين الظاهرين.

احسب $E(X)$ ، $\sigma^2(x)$

مثال:

محل لبيع الألبسة يربح في الأيام العادية عشرة دنانير في اليوم، وفي الأيام شديدة البرد يخسر خمسة دنانير، وفي أيام المواسم يربح مائة ديناراً.

فإذا علمت أن النسبة المئوية للأيام العادية وشديدة البرد والمواسم هي على الترتيب 60٪ ، 10٪، 30٪ . فإذا اختير أحد الأيام عشوائياً (Randomly). احسب توقع ربحه في ذلك اليوم.

الحل:

X	10	-5	100	Total
F(x)	0.6	0.1	0.3	1

Xf(x)	6	-0.5	30	35.5

توقع ربحه في ذلك اليوم = 35.5 دينار

نظرية: إذا كان X متغيراً عشوائياً فإن

$$E(ax+b) = aE(x) + b$$

حيث a,b عددان حقيقيان.

البرهان (Proof)

$$E(ax+b) = \sum_{x\in X(\Omega)} (ax+b) f(x)$$

$$= a \sum_{x\in X(\Omega)} Xf(x) + b \sum_{x\in X(\Omega)} f(x)$$

$$= aE(X) + b(1)$$

$$= aE(X) + b \square$$

مثال

إذا كان X متغيراً عشوائياً وكان E(X)= 5

احسب E (3X+1)

الحل:

$$E(3X+1) = 3E(x)+1$$

$$= 3(5) + 1$$

$$= 16$$

المتغير العشوائي المتصل (Continuous Random Variable)

تعريف: المتغير العشوائي المتصل X هو متغير عشوائي يكون مداه X(Ω) يحوي فترة .

ويدعى الاقتران f(x) اقتران كثافة احتمالية (Probability density function) للمتغير العشوائي X إذا كان.

1) $f(x) \geq 0 \quad \forall x \in X(\Omega)$

2) $\int_{X(\Omega)} f(x)\, dx = 1 \square$

مثال:

إذا كان X متغيراً عشوائياً مداه الفترة [0,4] وكان $f(x)=\frac{1}{8}x$. بين أن f(x) اقتران كثافة احتمالية (p.d.f).

الحل:

$f(x) \geq 0$ لكل x في الفترة [0,4]

$$\int_0^4 f(x)dx = \int_0^4 \frac{1}{8}xdx$$

$$= \frac{1}{16}X^2\Big]_0^4$$

$$= 1$$

$\therefore$ f(x) هو اقتران كثافة احتمالية (p.d.f)

مثال:

إذا كان $f(x) = Ae^{-x}$ فجد قيمة A التي تجعل f(x) اقتران كثافة احتمالية (p.d.f) للمتغير العشوائي X الذي يأخذ قيم في $(0,\infty)$.

الحل:

$$\int_0^\infty Ae^{-x}dx = 1$$

$$\Rightarrow \lim_{L\to\infty} \int_0^L Ae^{-x}dx = 1 \quad \text{(improper integral)}$$

$$\Rightarrow \lim_{L\to\infty} \left(-Ae^{-x}\Big]_0^L\right) = 1$$

$$\Rightarrow \lim_{L\to\infty} \left(-Ae^{-L} + A\right) = 1$$

$$\Rightarrow 0 + A = 1 \quad \Rightarrow \quad A = 1$$

تعريف:

إذا كان X متغير عشوائياً لا متصلاً مداه X(Ω) واقتران كثافته الاحتمالية f(x) فإن

1) توقع المتغير العشوائي X هو:

$$E(X) = \int_{x(\Omega)} Xf(x)dx$$

2) العزم (K) للمتغير العشوائي X هو

$$E(X^k) = \int_{\Omega} X^k f(x)dx$$

التباين (Variance) للمتغير العشوائي المتصل (X) هو

$$\sigma^2_{(x)} = E\left(X^2\right) - \left(E(x)\right)^2$$

والجذر التربيعي للتباين $\sigma = \sqrt{E\left(X^2\right) - \left(E(x)\right)^2}$ يسمى الانحراف المعياري.

مثال:

إذا كان X متغيراً عشوائياً اقتران كثافته الاحتمالية

$$f(x) = \frac{1}{8}x \ , \ x \in [0,4]$$

احسب:

1) E(x)
2) $E(x^2)$
3) σ^2 (x)

الحل:

1) $$E(X) = \int_0^4 x.\frac{1}{8}xdx$$

$$= \frac{1}{8}\int_0^4 x^2 dx$$

$$= \frac{1}{24}\int_0^4 x^3 \Big]_0^4$$

2) $$E(X^2) = \int_0^4 x^2 . \frac{1}{8} X dx$$

$$= \frac{1}{8}\int_0^4 x^3 dx$$

$$= \frac{1}{32}\int_0^4 x^4 \Big]_0^4$$

$$\therefore E(x^2) = \frac{256}{32}$$

$$= 8$$

3) $$\sigma^2(x) = E(X^2) - E^2(X)$$

$$= 8 - \left(\frac{8}{3}\right)^2$$

$$= 8 - \frac{64}{9}$$

$$= \frac{8}{9}$$

تمريــن:

للمتغير العشوائي X والذي يأخذ قيمة في الفترة $(0,\infty)$ واقتران كثافة الاحتمالية $f(x)=e^{-x}$. احسب: variance (x)

نظرية:

إذا كان X متغيراً عشوائياً متصلاً فإن

$$E(aX+b) = aE(X) + b \quad \forall\, a,b \in R$$

مثال:

إذا كان X متغيراً عشوائياً متصلاً حيث $E(X) = 3$

احسب $E(Y)$ حيث $Y = 5X-1$

الحل:

$$E(Y) = E(5X - 1)$$
$$= 5\,E(X) - 1$$
$$= (5)\,(3) - 1$$
$$= 14$$

تعريف:

إذا كان X متغيراً عشوائياً متصلاً، اقتران كثافة الاحتمالية (p.d.f) $f(x)$ فإن احتمال أي فترة من الفترات [a,b] ، [a,b) ، (a,b] ، (a,b)

يساوي

$$P(a \le X \le b) = \int_a^b f(x)dx$$

أي أن

$$P(a \le X \le b) = P(a < X \le b)$$
$$= P(a \le X < b)$$
$$= P(a < X < b)$$

ملاحظة:

احتمال أي مجموعة منتهية لمتغير عشوائي متصل يساوي صفراً.

مثال:

إذا كان X متغيراً عشوائياً متصلاً يأخذ قيمة في الفترة $[0,\infty)$ واقتران كثافته الاحتمالية هو

$$f(x) = e^{-x}$$

احسب:

1. $P(0 \le X < 1)$
2. $P(|X| < 2)$
3. $P\{0,1,2\}$

الحل:

1) $$P(0 \le X < 1) = \int_0^1 e^{-x} dx$$

$$= -e^{-x} \Big]_0^4$$

$$= -e^{-1} + 1$$

$$= 0.632$$

2) $$P(|X| < 2) = P(-2 < x < 2)$$

$$= P(0 \le x < 2)$$

$$= \int_0^2 e^{-x} dx$$

$$= 0.865$$

3) $P\{0,1,2\} = 0$

تمارين

1- إذا كـان $P(E_1)=\frac{2}{15}$ ، $P(E_2)=\frac{4}{15}$ وكـان E_1 ، E_2 حـادثين منفـصلين (disjoint event). جد $P(E_1 \cup E_2)$ ؟

2- إذا كـان E_1 ، E_2 حـادثين مـستقلين (Independent events) $P(E_1)=0.15$ ، $P(E_2)=0.4$ فجد

a) $P(E_1 \cap E_2)$ b) $P(E_1 \cup E_2)$

c) $P(\overline{E}_1 \cap \overline{E}_2)$ d) $P(\overline{E}_1 \cup E_2)$

3- كيس يحتوي (9) كرات سوداء، (6) كرات حمراء.سحب من الكيس كـرة واحـدة عشوائياً. احسب

a- احتمال أن تكون حمراء.

b- احتمال أن تكون سوداء.

c- احتمال أن تكون بيضاء.

4- كيس يحوي (9) كرات بيضاء، (11) كرة حمراء. سحب من الكيس كرتـان علـى التـوالي دون إرجـاع(without replacement) عـشوائياً. احـسب احتمـال أن تكون:

a- الكرتان حمراوتان.

b- الكرتان مختلفتان في اللون.

5- إذا كانـت E_1 ، E_2 ، E_3 حـوادث متباعـدة وشـاملة وكـان $P(E_1) = 0.2$ ، $P(E_2) = 0.3$ ، احسب $P(E_2)$؟

6- إذا كانت E_1 ، E_2 ، E_3 ، E_4 حوادث متباعدة وشاملة وكان

$P(E_1) = P(E_2) = 2P(E_3) = 2P(E_4)$

فأوجد $P(E_1), P(E_2), P(E_3), P(E_4)$

7- تقدم موظفان لأحد البنوك فإذا كان احتمال قبول الأول (0.7) واحتمـال قبـول الثاني (0.6) احسب

a- قبول الاثنين معا.

b- قبول الأول أو الثاني.

c- قبول الأول وعدم قبول الثاني.

d- عدم قبول الاثنين.

8- إذا كان ، $P(A) = 0.65$ ، $P(B)= 0.8$ ، $P(A \cap B) = 0.55$

a) $P(A/B)$ b) $P(B/A)$

c) $P(\overline{A}/B)$ d) $P(A \cap \overline{B})$

e) $P(A-B)$ f) $P(A/A \cup B)$

g) $P(\overline{A}/\overline{B})$

9- صندوق يحوي 9 كرات حمراء، 6 كـرات سـوداء، 5 كـرات بيـضاء، سـحب مـن الكيس كرتان على التوالي بشكل عشوائي مع الإرجاع (with replacement) مـا احتمال:

أ- أن تكون الكرة الأولى حمراء والثانية بيضاء.

ب- أن تكون الكرتان من نفس اللون.

ج- أن تكون الكرتان مختلفتان في اللون.

د- أن تكون إحدى الكرتين ليست سوداء.

10- أعد حل السؤال السابق إذا كان السحب دون إرجاع؟

11- صندوقان A، B يحوي A خمس كـرات حمـراء وسـبع كـرات سـوداء. ويحـوي B ثلاث كرات حمراء وأربع كرات سوداء. سـحبت كـرة مـن الـصندوقA عـشوائياً

ووضعت في الصندوق B، ثم سحب من الصندوق B كرتان عشوائياً دون إرجاع احسب احتمال أن تكون الكرتان المسحوبتان من الصندوق B:

a- حمراء ثم سوداء.

b- من نفس اللون.

12- إذا كان E_1 ، E_2 حادثين مستقلين، $P(E_1) = 2P(E_2)$ ، $P(E_1 \cup E_2) = \frac{5}{8}$ احسب $P(E_1)$

13- في رحلة لطائرة من عمان إلى جدة. إذا كان احتمال نفاد تذاكر الدرجة السياحية 80٪ واحتمال نفاد تذاكر الدرجة الأولى 70٪. واحتمال نفاد تذاكر الدرجتين معا 65٪. فإذا نفدت تذاكر الدرجة الأولى فما احتمال نفاد تذاكر الدرجة السياحية؟

14- تقدم طالب لامتحانين في الرياضيات والعلوم، فإذا كان احتمال نجاحه في الرياضيات 0.7، واحتمال نجاحه في العوم إذا نجح في الرياضيات 0.8، احسب احتمال نجاحه في المادتين معا؟

15- لتحديد النسل يصف الأطباء من خلال أحد مراكز الأمومة ثلاث وسائل A، B، C لمنع الحمل. فإذا كانت نسبة اللاتي تستخدمن هذه الوسائل هي 40٪، 35٪، 25٪ على الترتيب، وكانت نسبة الفشل في استخدام هذه الوسائل (كما حددها الأطباء) هي 5٪، 2٪، 1٪ على الترتيب. اختيرت إحدى النساء عشوائياً وكانت تستخدم إحدى هذه الوسائل . احسب احتمال أن تكون قد استخدمت الوسيلة B إذا علمت أنها حامل؟

16- متحف 60٪ من رواد عرب والباقي أجانب. فإذا كانت نسبة الرواد الذكور من العرب 95٪، ونسبة الرواد الذكور من الأجانب 20٪، فإذا اختير أحد الرواد وكان عربياً فما احتمال أن تكون أنثى؟

17- سائق تكسي يحمل في جعبته ثلاثة دنانير قطع معدنية فإذا كان الدينار الأول من فئة الخمسة قروش والدينار الثاني من فئة العشرة قروش أما الثالث فكانت من فئة الربع دينار. إذا سحب السائق من الجعبة قطعتان نقديتان معا ودل المتغير العشوائي X على قيمة القطعتين، احسب:

a- مدى X.

b- التوزيع الاحتمالي للمتغير العشوائي.

c- E(x).

اعتبر جميع الفئات النقدية لها نفس فرصة الاختيار.

18- إذا كان جدول التوزيع الاحتمالي للمتغير العشوائي x.

X	1	2	3
$P(X=x)$	a	2a	b

وكان E(x)= 2.2 احسب قيمة كل من a , b؟

19- إذا كان توقع ربح شخص في مسابقتين له نفس القيمة، وكانت قيمة الجائزة الأولى (250) ديناراً واحتمال الحصول عليها $\left(\frac{2}{5}\right)$ مما مقدار الجائزة الثانية إذا كان احتمال الحصول عليها $\frac{4}{5}$؟

20- إذا كان X متغيراً عشوائياً يأخذ قيمه في المجموعة {0,1,2} وكان اقتران احتماله هو $f(x) = \frac{1}{5}x^2$. احسب E(x)؟

21- إذا كان X متغيراً عشوائياً يأخذ قيمه في الفترة [0,2] وكان اقتران كثافته الاحتمالية f(x)=bx-2. جد E(X)؟

22- إذا كان التوزيع الاحتمالي للمتغير العشوائي X هو

{(-4,a), (3,4a), (0,7a)} فجد قيمة a؟

23- صف به (10) أولاد (5) بنات. إذا اختير عشوائياً (Randomly) ثلاثة طلاب على التوالي. احسب احتمال:

a- أن يكون الأول والثاني ولدين والثالثة بنتاً.

b- أن يكون الأول والثالث ولدين والثانية بنتاً.

c- أن يكون الأول والثالث من الجنس نفسه والثاني من الجنس الآخر.

24- الصندوق A يحوي (5) كرات حمراء، (3) بيضاء، (8) زرقاء. أما الصندوق B فيحوي (3) كرات حمراء، (5) بيضاء. إذا ألقي حجر نرد منتظمة فإننا نسحب كرة من صندوق B إذا ظهر الوجه (3) أو (6). وغير ذلك نسحب كرة من الصندوق A.

a- احسب احتمال أن تكون الكرة المسحوبة حمراء.

b- إذا كانت الكرة المسحوبة حمراء فما احتمال أن تكون من الصندوق A.

25- صندوق يحوي قطعتي نقد إحدهما منتظمة والأخرى على وجهيها صور. فإذا سُحب عشوائياً قطعة منها وألقيت، فإنه إذا ظهر صورة نلقي القطعة الأخرى، أما إذا ظهر كتابة فنلقي نفس القطعة مرة أخرى. احسب احتمال :

a- ظهور صورة في الرمية الثانية.

b- إذا ظهر في الرمية الثانية صورة. فما احتمال أن تكون قد ظهرت صورة في الرمية الأولى.

26- في تجربة رمي حجري النرد، إذا دلّ المتغير العشوائي X على أنه العدد الأكبر للوجهين الظاهرين، أو أحدهما إذا كانا متساويين. احسب توقع X.

27- عند إلقاء حجر نرد مرة واحدة يربح شخص مبلغاً من الدنانير مساوياً لعدد نقط الوجه الظاهر إذا كان الوجه الظاهر عدد أولي، ويخسر دنانير مساوية لعدد النقط الظاهرة على وجه حجر النرد إذا ظهر عدد غير أولي. احسب توقع ربح هذا الشخص.

28- إذا كـــــان X مـــــتغيراً عـــــشوائياً، اقـــتران كثافتـــــه الاحتماليـــــة

$f(x) = \frac{1}{6}x + a \ , \ x \in [0,3]$ احسب

a- قيمةa　　　b -P(1 ≤X ≤2)

c -E(X)　　　d -P(X^2<4)

e -P{1,2}　　　f -E(16X)

29- إذا كان الفضاء العيني لتجربة عشوائية $\Omega = \{a_1, a_2, a_3, a_4\}$، p اقتران احتمال

بحيث $P(\{a_2, a_3\}) = \frac{2}{3}$ ، $P(\{a_2, a_4\}) = \frac{1}{2}$ ، $P(\{a_2\}) = \frac{1}{3}$

احسب $P(\{a_1\})$

30- ستة رجال وزوجاتهم في غرفة، اختير منهم شخصين عـشوائياً. احـسب احتمـال أن يكون الشخصين:

a- زوج وزوجته　　b- مختلفين في الجنس

31- في السؤال السابق، إذا اختير أربعة أشخاص، احسب احتمال أن يكون

a- هؤلاء عائلتين "كل عائلة مكونة من زوج وزوجته".

b- لا يوجد أية عائلة بينهم.

c- بينهم عائلة واحدة فقط.

32- إذا كان A ، B مستقلين بحيث أن P(A)=k ، P(B)= k+0.2 ، $P(A \cap B) = 0.15$

a- احسب قيمة k.

b- $P(A \cup B)$

c- $P(\overline{A}/\overline{B})$

6

الوحدة السادسة

التوزيعات الاحتمالية

Probability Distributions

التوزيعات الاحتمالية

Probability Distributions

مقدمة

يسمى التوزيع الاحتمالي الذي متغير العشوائي منفصلاً توزيع احتمالي منفصل أما التوزيع الذي متغيره العشوائي متصلاً فيسمى توزيعاً احتمالياً متصلاً.

وستتعرف في هذه الوحدة على توزيعات احتمالية مهمة منها ما هو منفصل ومنها ما هو متصل.

التوزيعات الاحتمالية المنفصلة Discrete Probability distributions

1- توزيع ذات الحدين Binomial distribution

تعالج نظرية ذات الحدين ذلك النوع من التجارب التي تتكرر عدد محدود من المرات وتكون نتيجتها في المرة الواحدة أما نجاح أو فشل.

نظرية:

إذا أجريت تجربة (n) مرة، وكان احتمال نجاحها في المرة الواحدة هو (p)، ودل المتغير العشوائي x على عدد مرات النجاح فإن احتمال نجاح التجربة في x مرة هو

1. $P(x)=\binom{n}{x}p^{x}(1-p)^{n-x} \quad , \; x=0,1,2,\ldots n$
2. $E(X)=np$ توقع المتغير العشوائي X هو
3. $\sigma^{2}(X)=npq$ تباين المتغير العشوائي X □

مثال:

إذا كان احتمال نجاح عملية جراحية هو (0.9) فإذا أجريت العملية لعشرة مرضى احسب ما يلي:

a- احتمال نجاح العملية لسبعة مرضى.

b- احتمال نجاح العملية لجميع المرضى.

c- احتمال نجاح العملية لثمانية مرضى على الأقل.

d- احتمال نجاح العملية لثمانية مرضى على الأكثر.

e- توقع عدد المرضى الذين سيجرون العملية بنجاح.

f- تباين عدد المرضى الذين سيجرون العملية.

الحل:

n=10 , p=0.9

$$P(x)=\binom{10}{x}(0.9)^{x}(0.1)^{10-x}, x=0,1,2,...,10$$

a) $$P(7)=\binom{10}{7}(0.9)^{7}(0.1)^{3}$$

b) $$P(10)=\binom{10}{10}(0.9)^{10}(0.1)^{0}$$
$$=(0.9)^{10}$$

c) $$P(X\geq 8)=[P(8)+P(9)+P(10)]$$
$$=\sum_{x=8}^{10}\binom{10}{x}(0.9)^{x}(0.1)^{10-x}$$

d) $$P(X\leq 8)=P(0)+P(1)+...+P(8)$$
$$=\sum_{x=0}^{8}\binom{10}{x}(0.9)^{x}(0.1)^{10-x}$$

حل آخر:

$$P(x\leq 8)=1-[P(x=9)+P(x=10)]\square$$

$$= 1 - \sum_{x=9}^{10} \binom{10}{x} (0.9)^x (0.1)^{10-x}$$

e) E(X)= np□
= (10) (0.9)
= 9

f) $\sigma^2(f) = npq = (10)(0.9)(0.1)$
= 0.9

مثال:

ألقي حجر نرد إحدى وخمسون مرة، إذا كان المتغير العشوائي X هو عدد مرات الحصول على عدد يقبل القسمة على (3). احسب

a) احتمال نجاح التجربة (الحصول على عدد يقبل القسمة على (3)) عدد من المرات لا يقل عن (20) ولا يزيد عن (30).

b) احتمال عدم نجاح التجربة.

E(X) (c

σ(X) (d

الحل:

$$n = 51, P = \frac{2}{6} = \frac{1}{3} \Rightarrow P(x) = \binom{51}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{51-x}$$ □

a) $P(20 \le x \le 30) = \sum_{x=20}^{30} \binom{51}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{51-x}$

b) $P(0) = \binom{51}{0} \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^{51} = \left(\frac{2}{3}\right)^{51}$

□

c) $E(X) = 51 \times \frac{1}{3} = 17$

d) $\sigma(X) = \sqrt{npq} = \sqrt{(51)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)} = \sqrt{11.3} = 3.36$

2- توزيع بواسون Poisson Distribution

يهتم توزيع بواسون في تلك التجارب التي تحدث خلال فترة زمانية أو مكانية محددة كدراسة عدد المكالمات التي تصل مقسم ما خلال ساعات الدوام. أو دراسة عدد حوادث السير عند تقاطع معين خلال أسبوع معين. فإذا كان معدل النجاح في فترة زمانية (مكانية) محددة هو (λ) فيكون احتمال بواسون معطى بالعلاقة

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0,1,2,\ldots$$

ويكون توقع بواسون $E(x) = \lambda$

وتباين بواسون $\sigma^2 = \lambda$

ويمكن تقريب توزيع ذات الحديث إلى توزيع بواسون بوضع $\lambda = np$ إذا كان (n) كبيرة جداً و (p) صغيرة جداً.

مثال:

إذا كان متوسط عدد الأيام التي تمطر فيها في شهر شباط هي ثلاثة أيام في الأسبوع. فما احتمال أن تمطر خمسة أيام في الأسبوع في ذلك الشهر.

الحل:

$\lambda = np$

x= 5

λ=3

$$\therefore P(5) = \frac{e^{-3}5^{3}}{5!} = \frac{(0.05)(729)}{120} = 0.3$$

التوزيعات الاحتمالية المتصلة

Continuous Probability distributions

1- التوزيع الطبيعي: Normal distribution

قبل البدء بموضوع التوزيع الطبيعي لنتعرف على مفهوم العلامة المعيارية.

العلامات المعيارية Standard mark

تعريف: إذا كان لدينا مجموعة من المفردات وسطها الحسابي $\overline{X}$ وانحرافها المعياري σ وإذا كانت X مفردة ما "تسمى العلامة الخام Row mark" فإن العلامة المعياريةZ المناظرة لها هي:

$$Z = \frac{X - \overline{X}}{\sigma}$$

وتستخدم العلامات المعيارية لمقارنة علامتين من توزيعين مختلفين، فتكون المقارنة أكثر عدالة.

مثال:

إذا كان الوسط الحسابي لمجموعة من المفردات (50) والانحراف المعياري (10) فأوجد:

1- العلامة المعيارية المناظرة للعلامة الخام60.
2- العلامة المعيارية المناظرة للعلامة الخام45.
3- العلامة المعيارية المناظرة للوسط الحسابي.
4- العلامة الخام المناظرة للعلامة المعيارية 1.5 .

الحل:

$$\overline{X} = 50 \ , \ \sigma = 10$$

1) $$Z = \frac{X - \overline{X}}{\sigma}$$

2) $$= \frac{60 - 50}{10} = 1$$

$$= \frac{45-50}{10} = -0.5$$

3) $$Z = \frac{X-\overline{X}}{\sigma}$$

$$= \frac{50-50}{10} = 0$$

4) $$Z = \frac{X-\overline{X}}{\sigma}$$

$$\Rightarrow 1.5 = \frac{x-50}{10}$$

$$\Rightarrow 15 = x-50 \Rightarrow x = 65$$

مثال:

طالب في شعبة A علامته في مادة الإحصاء 60، وطالب آخر في شعبة B علامته في الإحصاء 70، فإذا علمت أن الوسط الحسابي لعلامات طلبة شعبة A في الإحصاء (65) والانحراف المعياري لها (5) أما طلبة شعبة B فالوسط الحسابي لعلاماتهم في الإحصاء (85) والانحراف المعياري لها (10) فأي الطالبين تحصيله أفضل في الإحصاء هل هو طالب شعبة A. أم طالب شعبة B.

الحل:

شعبة A: $x = 60$, $\sigma = 5$, $\bar{x} = 65$

إذن فالعلامة المعيارية للطالب الموجود في شعبة A هي:

$$Z_A = \frac{60-65}{5}$$

$$= -1$$

شعبة B: $x = 70$, $\sigma = 10$, $\bar{x} = 85$

$$Z_B = \frac{70-85}{10}$$

$$= -1.5$$

النتيجة : $-1 > -1.5$ أي أن العلامة المعيارية للطالب الموجود في شعبة B أصغر من نظيرتها للطالب الموجود في شعبة A. وبالتالي تحصيل الطالب الموجود في شعبة A أفضل في الإحصاء من تحصيل نظيره الموجود في شعبة B.

مثال:

أحمد وعثمان طالبان في الصف الأول الثانوي العلمي، فإذا كانت علامة أحمد في الرياضيات هي (72) والعلامة المعيارية لها هي (1.5)، وكانت علامة عثمان في نفس المادة (80) والعلامة المعيارية المقابلة لها (2.5). احسب الوسط الحسابي والانحراف المعياري لطلاب الصف في مادة الرياضيات؟

الحل:

أحمد: $X_1=72 \ , \ Z_1=1.5$

عثمان: $X_2=80 \ , \ Z_2=2.5$

$$Z_1 = \frac{x_1 - \bar{x}}{\sigma} \Rightarrow 1.5 = \frac{72 - \bar{x}}{\sigma} \Rightarrow 1.5\sigma + \bar{x} = 72 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$Z_2 = \frac{x_2 - \bar{x}}{\sigma} \Rightarrow 2.5 = \frac{80 - \bar{x}}{\sigma} \Rightarrow 2.5\sigma + \bar{x} = 80 \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

بحل المعادلتين (1)، (2) ينتج أن

الوسط الحسابي $\bar{x} = 60$

والانحراف المعياري $\sigma = 8$

التوزيع الطبيعي Normal Distribution

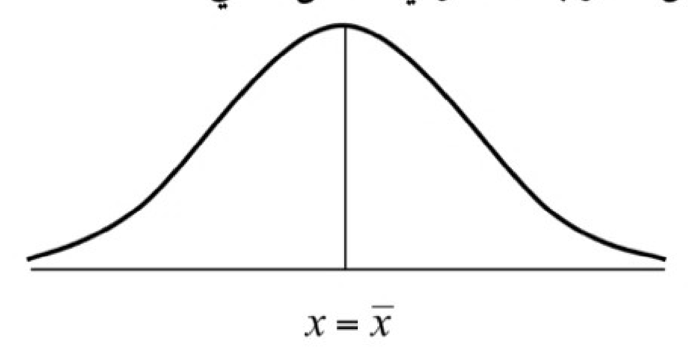
التوزيع الطبيعي هو ذلك التوزيع الذي يكون متغيرة العشوائي متصل واقتران كثافته الاحتمالية هو:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

العدد النيبيري = e

النسبة التقريبة= π

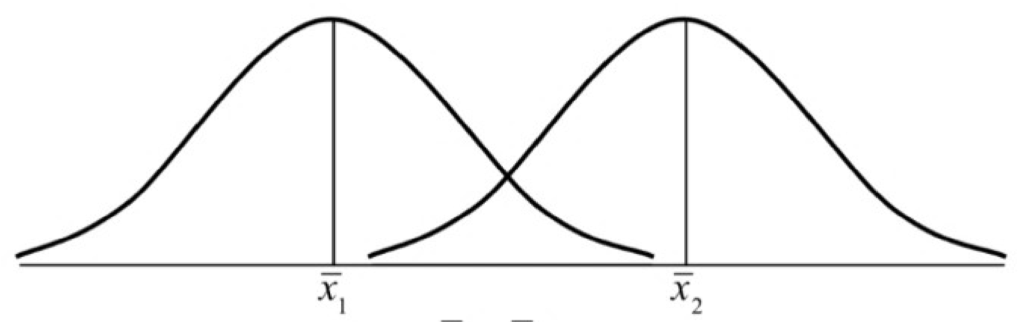
ويكون الوسط الحسابي للمتغير العشوائي هو $\bar{x}$ وانحرافه المعياري σ ومنحناه يأخذ شكل الناقوس المقلوب كما هو في الشكل التالي:



$x = \bar{x}$

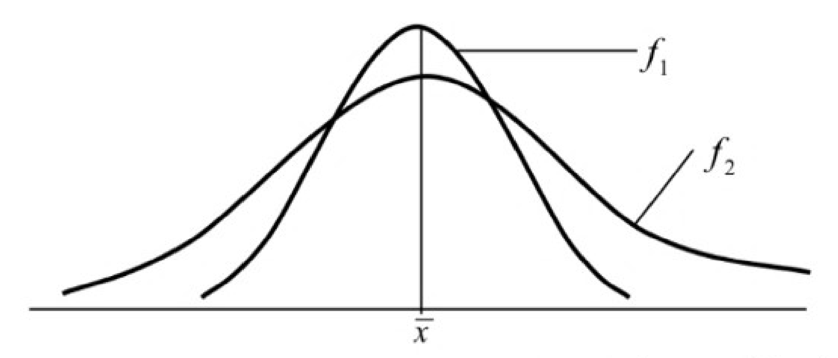
خصائص التوزيع الطبيعي:

1- منحناه يأخذ شكل الناقوس المقلوب ويمتد من طرفيه إلى -∞ ، ∞.
2- المساحة المحصورة بين منحنى اقتران كثافته الاحتمالية ومحور السينات تساوي وحدة مربعة واحدة.
3- يكون المنحنى متماثلاً حول المستقيم $x = \bar{x}$.
4- أحادي المنوال.
5- الوسط الحسابي = الوسيط = المنوال.
6- بتغير $\bar{x}$ وثباتσ يتحرك المنحنى أفقياً يميناً أو يساراً كما هو موضح في الشكل التالي:



$\bar{x}_1 < \bar{x}_2$

7- كلما ازدادت σ وبقيت $\bar{x}$ ثابتة فإن المنحنى يبتعد أكثر عن الوسط الحسابي من الجهتين. كما هو موضح في الشكل التالي:



f_1 يرتبط بالانحراف المعياري σ_1

f_2 يرتبط بالانحراف المعياري σ_2 وتكون $\sigma_1<\sigma_2$

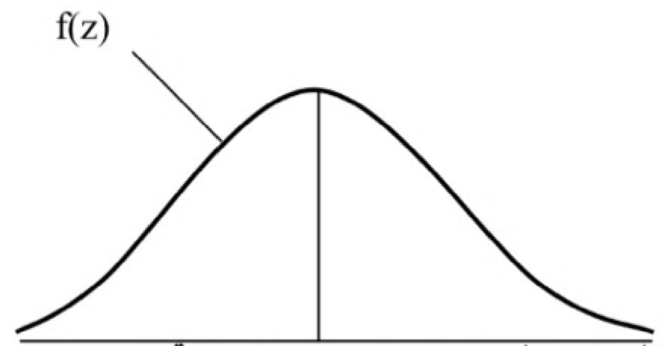
8- التوزيع الطبيعي الذي وسطه الحسابي (صفراً) وانحرافه المعياري (1) يسمى توزيعاً طبيعياً معيارياً.

9- يمكن تحويل أي توزيع طبيعي وسطه الحسابي $\bar{x}$ وانحرافه المعياري σ إلى توزيع طبيعي معياري باستخدام العلاقة $z = \dfrac{x-\bar{x}}{\sigma}$.

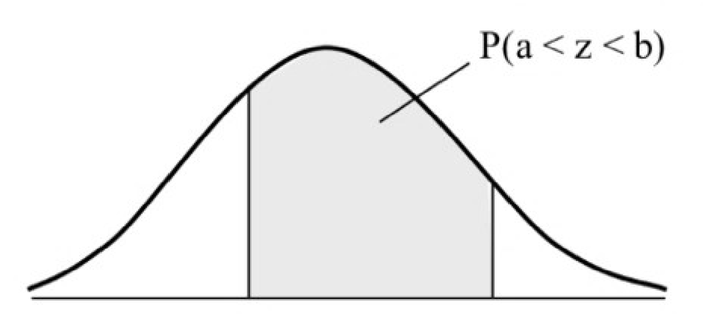
التوزيع الطبيعي المعياري Standard Normal Distribution

التوزيع الطبيعي المعياري هو التوزيع الطبيعي الذي وسطه الحسابي (صفر) وانحرافه المعياري (1) ويكون اقتران كثافته الاحتمالية هو:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad -\infty < Z < \infty$$



سنرمز لنسبة المساحة الواقعة بين z=a، z=b وتحت منحنى التوزيع الطبيعي وفوق محور x بالرمز $P(a<z<b)$.



ملاحظة:

$$P(a \le z \le b) = P(a<z<b)$$
$$= P(a < z \le b)$$
$$= P(a \le z < b)$$

وذلك لأن المساحة الواقعة فوق نقطة تساوي صفراً.

أما عن كيفية إيجاد نسبة المساحة الواقعة تحت المنحنى الطبيعي المعياري وأي قيمتين فنستخدم جداول خاصة ملحقة بنهاية الكتاب. تعطي نسبة المساحة الواقعة بين z=0 وأية قيمة موجبة. وباستخدام خصائص المنحنى الطبيعي يمكننا إيجاد نسبة المساحة المطلوبة.

مثال:

أوجد:

1) $P(0<z<1)$

2) $P(0<z<1.5)$

3) $P(0<z<1.25)$

الحل:

جميع هذه النسب نجدها من جدول التوزيع الطبيعي المعيـاري الموجـود في الملحـق رقم (2) مباشرة.

1) $P(0<z<1) = 0.3413$

2) $P(0<z<1.5) = 0.4332$

3) $P(0<z<1.25) = 0.3944$

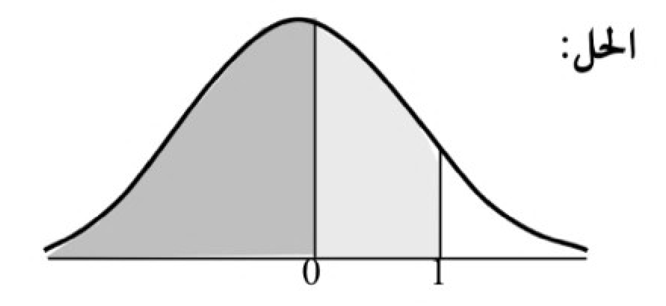
مثال:

أوجد:

1) $P(z<1)$

2) $P(z<1.5)$

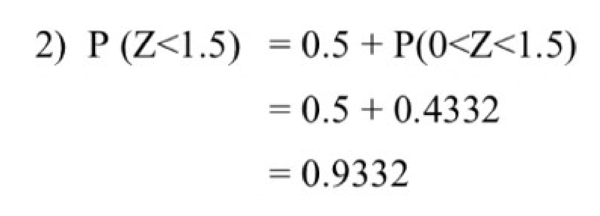
الحل:



1) $P(z<1) = 0.5 + P(0<z<1)$

$= 0.5 + 0.3413$

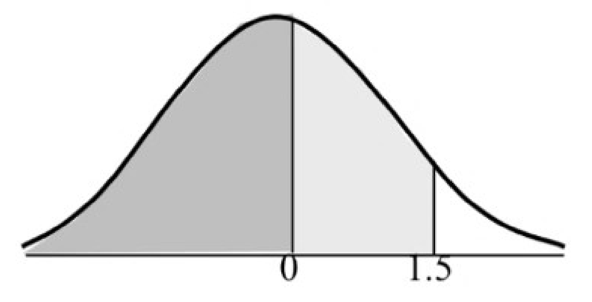
$= 0.8413$



2) $P(Z<1.5) = 0.5 + P(0<Z<1.5)$

$= 0.5 + 0.4332$

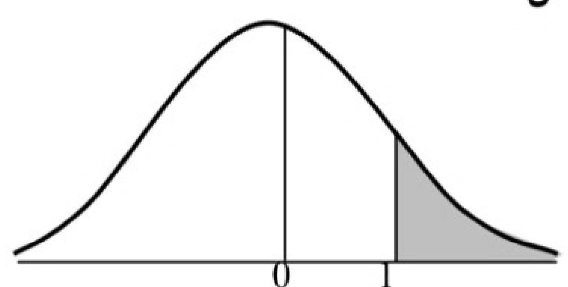
$= 0.9332$



مثال :

أوجد 2) $P(Z>1.5)$ 1) $P(Z>1)$

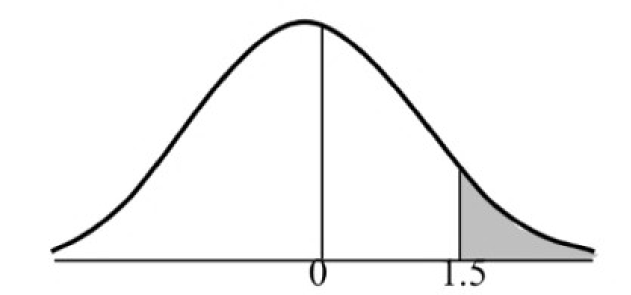
الحل:



1) $P(z>1) = 0.5 - P(0<Z<1)$

$= 0.5 - 0.3413$

$= 0.1587$



2) $P(z>1.5) = 0.5 - P(0<z<1.5)$

$= 0.5 + 0.4332$

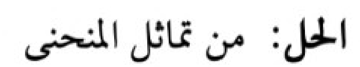
$= 0.0668$

مثال:

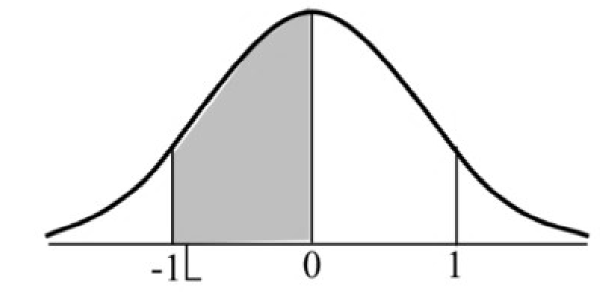
أوجد :

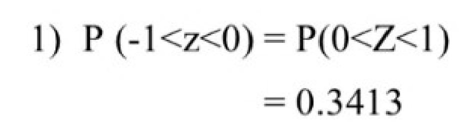
1) $p(-1<z<0)$

2) $p(-2<z<0)$ □



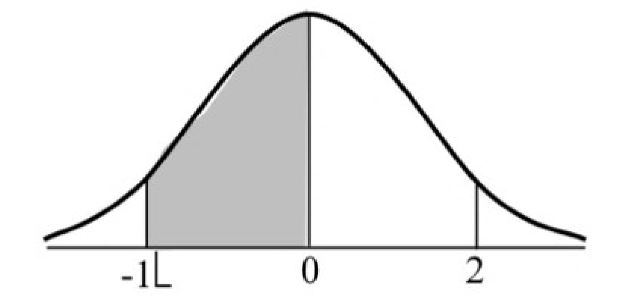
الحل: من تماثل المنحنى

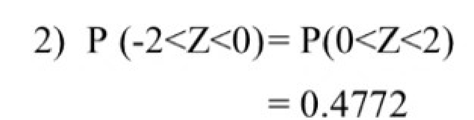




1) $P(-1<z<0) = P(0<Z<1)$

$= 0.3413$





2) $P(-2<Z<0) = P(0<Z<2)$

$= 0.4772$

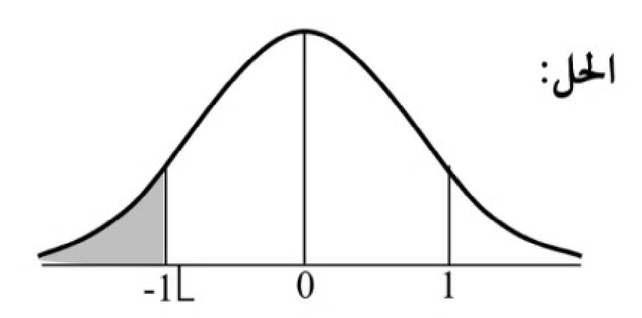
مثال:

أوجد :

1) $p(z<-1)$

2) $p(z<-2)$

الحل:



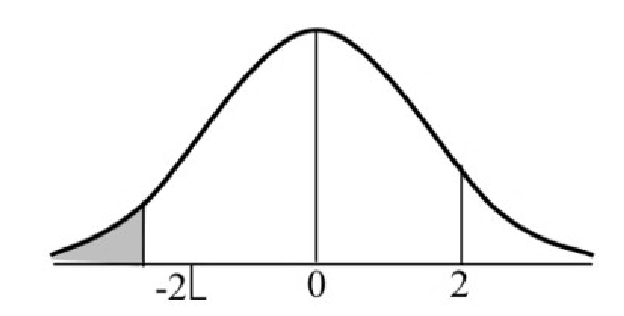
1) $P(z<-1) = P(z>1)$

$= 0.5 - p(0<z<1)$

$= 0.5 - 0.3413$

$= 0.1587$

□



2) $P(Z<-2) = P(Z>2)$

$= 0.5 - p(0<Z<2)$

$= 0.5 - 0.4772$

$= 0.0228$

مثال:

□ أوجد

1- $P(1<z<2)$

2- $P(0.5<z<1.5)$

3- $P(-2<z<-1.5)$

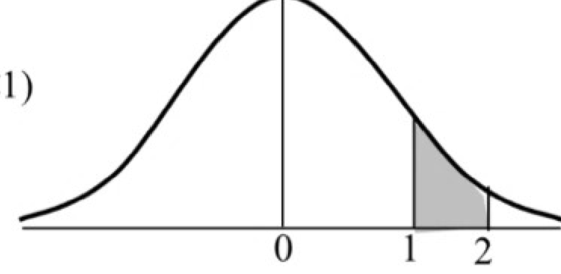
4- $P(-1<z<-0.5)$

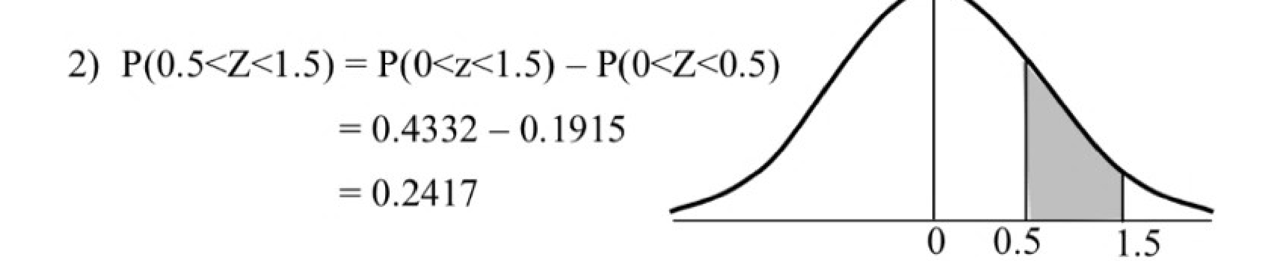
الحل:

1) $P(1<Z<2) = P(0<Z<2) - P(0<Z<1)$

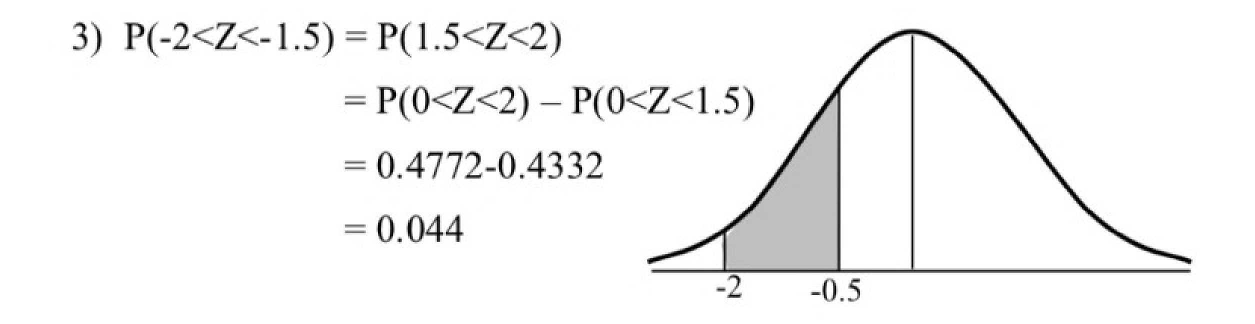
$= 0.4772 - 0.3413$

$= 0.1359$

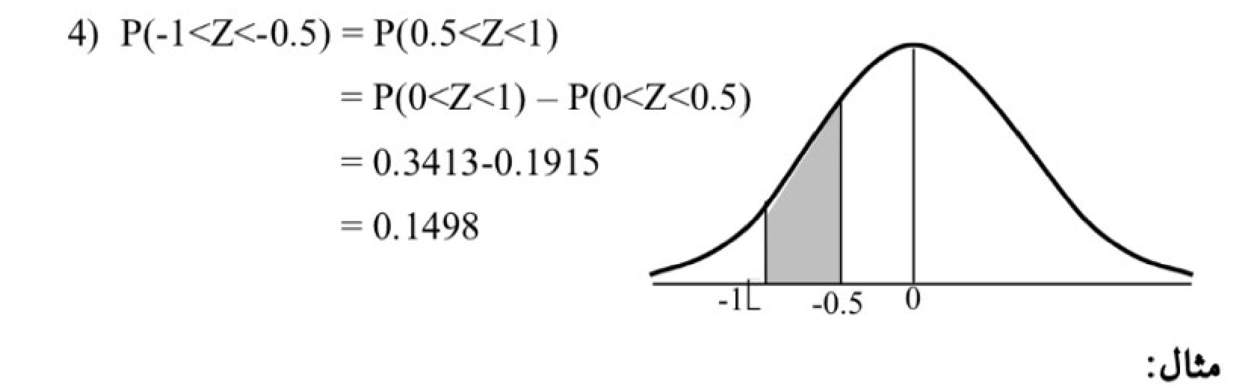




$$2)\ P(0.5<Z<1.5) = P(0<z<1.5) - P(0<Z<0.5)$$
$$= 0.4332 - 0.1915$$
$$= 0.2417$$



$$3)\ P(-2<Z<-1.5) = P(1.5<Z<2)$$
$$= P(0<Z<2) - P(0<Z<1.5)$$
$$= 0.4772-0.4332$$
$$= 0.044$$



$$4)\ P(-1<Z<-0.5) = P(0.5<Z<1)$$
$$= P(0<Z<1) - P(0<Z<0.5)$$
$$= 0.3413-0.1915$$
$$= 0.1498$$

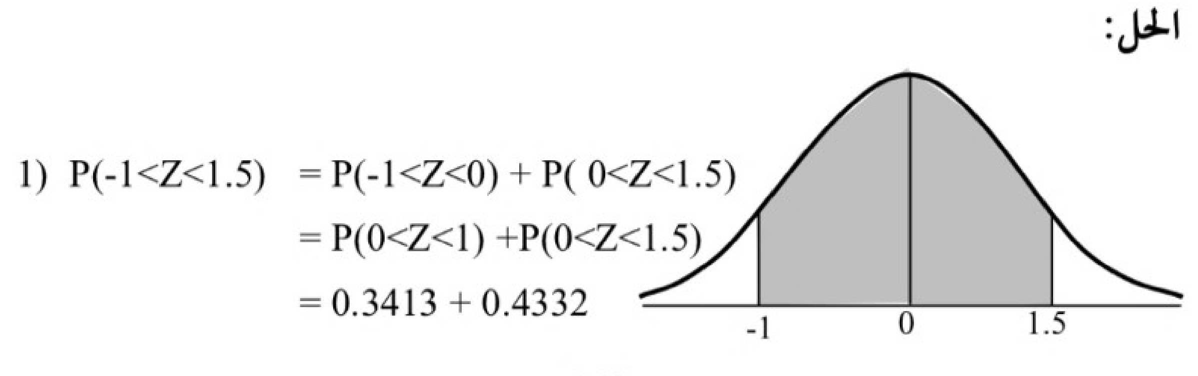
مثال:

أوجد:

1) $P(-1<Z<1.5)$

2) $P(|Z|<2)$

الحل:



$$1)\ P(-1<Z<1.5) = P(-1<Z<0) + P(0<Z<1.5)$$
$$= P(0<Z<1) + P(0<Z<1.5)$$
$$= 0.3413 + 0.4332$$

$= 0.7745$

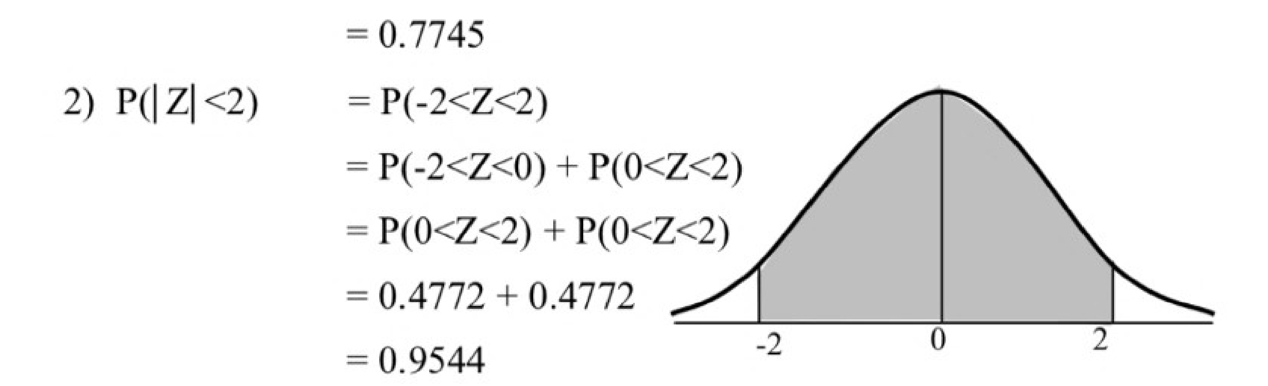
2) $P(|Z|<2) = P(-2<Z<2)$

$= P(-2<Z<0) + P(0<Z<2)$

$= P(0<Z<2) + P(0<Z<2)$

$= 0.4772 + 0.4772$

$= 0.9544$



مثال:

أوجد قيمة a بحيث أن $P(0 < Z < a) = 0.4251$

الحل:

بعد البحث في جدول التوزيع الطبيعي المعياري ضمن المساحات نجد $a=1.44$.

مثال:

أوجد قيمة a بحيث أن $P(Z<a) = 0.8413$

الحل:

قيمة a التي تحقق العلاقة $P(Z<a)= 0.8413$ هي نفس قيمة a التي تحقق العلاقة $P(0<Z<a)= 0.3413$ "لماذا؟". ومن جدول التوزيع الطبيعي المعياري نجد أن $a=1$.

مثال:

أوجد a بحيث أن $P(Z>a) = 0.33$

الحل:

قيمة a التي تحقق العلاقة $P(Z>a)= 0.33$ هي نفس قيمة a التي تحقق العلاقة $P(0<Z<a)=0.17$ ومن جدول التوزيع الطبيعي المعياري نجد أن

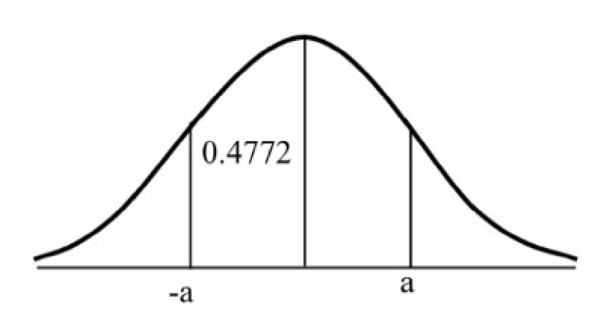
$a= 0.44$ □



مثال:

أوجد قيمة a بحيث $P(a<Z<0)= 0.4772$

الحل:



a التي تحقق العلاقـة $P(-a<Z<0)=$

0.4772 هي نفس a الـتي تحقـق العلاقـة $P(-a<Z<0)= 0.4772$

ومـــن جـــدول التوزيـــع الطبيعـــي المعياري نجد أن $-a=2$ ومنها $a=-2$

مسائل عملية على التوزيع الطبيعي

نعيد التذكير في أنه لتحويل أي توزيع طبيعي وسطه الحسابي $\bar{x}$ وانحرافه المعيـاري σ إلى توزيع طبيعي معياري، فإننا نستخدم العلاقة

$$Z = \frac{x - \bar{x}}{\sigma}$$

مثال:

تتخذ أطوال (1000) طالب توزيعا طبيعياً وسـطه الحـسابي (160cm) وانحرافـه المعياري (10cm).

أوجد:

1- نسبة الطلبة الذين تقل أطوالهم عن 170cm.

2- النسبة المئوية للطلبة الذين تزيد أطوالهم عن 180cm.

3- النسبة المئوية للطلبة الذين تتراوح أطوالهم بين165cm, 175cm.

4- عدد الطلبة الذين تزيد أطوالهم عن 175cm.

الحل:

$$n = 100 \quad x = 160cm \;, \quad \sigma = 10cm\ s$$

1- نسبة الطلبة الذين تقل أطوالهم عن 170 cm.

$$P(X<170)=P\left(\frac{x-\bar{x}}{\sigma}<\frac{170-\bar{x}}{\sigma}\right)$$

لكن
$$Z=\frac{x-\bar{x}}{\sigma}$$

إذاً: نسبة الطلبة الذين تقل أطوالهم عن 170cm هي:

$$P\left(Z<\frac{170-160}{10}\right)$$

$= P(Z<1)$

$= 0.5 + p(0<Z<1)= 0.8413$

إذاً: نسبة الطلبة الذين تقل أطوالهم عن 170cm تساوي(0.8413)

2- نسبة الطلبة الذين تزيد أطوالهم عن 180cm هي:

$P(x >180)$

$$=P\left(Z>\frac{180-160}{10}\right)$$

$= P(Z>2)$

$= 0.5 – 0.4772$

$= 0.0228$

إذاً: فالنسبة المئوية للطلبة الذين تزيد أطوالهم عن 180cm تساوي

$0.0228 \times 100\%$

$= 2.28\%$ □

3- نسبة الطلبة الذين تتراوح أطوالهم بين 165cm, 175cm هي

$$p(165<x<175)=P\left(\frac{165-160}{10}<Z<\frac{175-160}{10}\right)$$

$= P(0.5 < Z < 1.5)$

$= P(0<Z<1.5) – P(0<Z<0.5)$

$$= 0.4332 - 0.1915$$
$$= 0.2417$$

إذن النسبة المئوية للطلبة الذين تتراوح أطوالهم بين105cm , 175cm تساوي 24.17%

4- لإيجاد عدد الطلبة الذين تزيد أطوالهم عن 175cm، نجد نسبتهم ونـضربها في عدد الطلبة الكلي، أي أن عدد الطلبة المطلوب = نسبتهم × عدد الطلبة الكلي.

نسبة الطلبة الذين تزيد أطوالهم عن 175 cm هي:

$$P(X > 175) = P\left(Z > \frac{175 - 160}{10}\right)$$
$$= P\ (Z>1.5)$$
$$= 0.5 - P(0<Z<1.5)$$
$$= 0.5 - 0.4332$$
$$= 0.0668 \square$$

إذن فعدد الطلبة الذين تزيد أطوالهم عن 175cm يساوي:

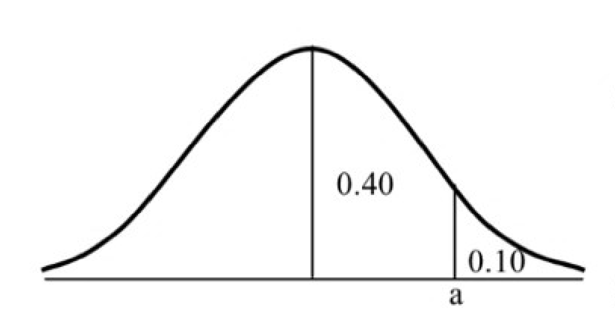
$$0.0668 \times 1000$$
$$= 66.8$$
$$\approx 67 \square$$

مثال:

إذا كان علامات طلبة الثانوية العامة في أحد الأعوام، تتخذ توزيعا طبيعياً وسـطه الحسابي (65) وانحرافه المعياري (15). فإذا علمت أن الجامعات الرسمية قبلـت علـى المعدل التنافسي 10٪ من هؤلاء الطلبة. فما هو أدنى معدل قبل في الجامعات الرسميـة حسب المعدل التنافسي؟

الحل:

$\bar{x} = 65 \;,\; \sigma = 15$



لإيجاد أدنى علامة معيارية قبلت في الجامعات الرسمية لنفرضها a فيكون

$P(Z<a)= 0.10$

وقيمة a التي تحقق العلاقة السابقة هي نفس قيمة a

التي تحقق العلاقة $P(0<Z<a)= 0.4$

الآن نبحث في جدول التوزيع الطبيعي عن المساحة (0.4000) فلا نجدها لذلك نأخذ أقرب قيمة لها وهي (0.3997) وبالتالي فإن a=1.28 وهذه هي أدنى علامة معيارية قبلت في الجامعات الرسمية ولتحويلها إلى علامة خام نستخدم العلاقة

$$Z = \frac{x-\bar{x}}{\sigma} \Rightarrow 1.28 = \frac{x-65}{15}$$

$$\Rightarrow x\text{-}65 = 1.28 \times 15$$

$$\Rightarrow x\text{-}65 = 19.2$$

$$\Rightarrow x= 65+19.2 = 84.2$$

(أي أدنى علامة قبلت في الجامعات الرسمية حسب المعدل التنافسي هي 84.2)

مثال:

بالاعتماد على المثال السابق، احسب

a- P_{40}

b- Q_3 الربيع الثالث

الحل:

a) $\bar{X} = 65 \;,\; \sigma = 15$

P_{40} هو العلامة التي يقل عنها أو يساويها 40% من العلامات ويقابل معياريا القيمة a والتي تحقق $P(Z<a)= 40\%$

$= 0.4000$

a التي تحقق العلاقة $P(Z<a)=0.4000$ هي نفس a التي تحقق العلاقة

$$P(a<Z<0) = 0.1000$$
$$\Rightarrow a= -0.25$$
$$\therefore -0.25 = \frac{P40-65}{15}$$
$$\Rightarrow P_{40}= 61.25$$

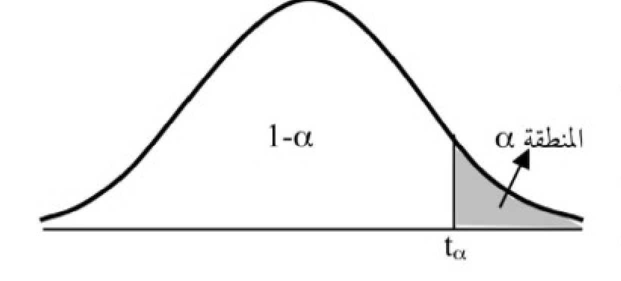
b) الربيع الثالث $P_{75}= 75.05$ "تأكد من ذلك"

2- توزيـــــع t: (t Distribution)

يشبه توزيع t التوزيع الطبيعي حيث يكون متماثل حول الوسط الحسابي الـذي يساوي صفر (t=0) حيث يكون اقتران الكثافة الاحتمالية له هو:

$$f(t) = c\left(1+\frac{t^2}{\gamma}\right)^{-\gamma+\frac{1}{2}} \qquad -\infty < t < \infty$$

ويسمى أيضاً توزيع ستيودانت (Student's distribution) حيث تمثل γ درجات الحرية df (degree of freedom)



كما نلاحظ هنا فإن شكل توزيع t يشبه التوزيع الطبيعي وله نفس الخصائص حيث يمتد من طرفيه من ∞- إلى ∞ ومتماثل حول محور الوسط. $P(t>t_\alpha)= \alpha$

والمساحة تحته = وحدة مربعة واحدة.

ولكنه يختلف عن التوزيع الطبيعي في كون قيمة تعتمد على درجات الحرية (df).

ولإيجاد قيمة t نجد درجات الحرية وهي (df=n-1) ونجدها من الجدول المخصص لها بحيث تكتب على الصورة $t[df, \alpha]$.

مثال: احسب قيمة $t[15, 0.05]$

الحل: نجد قيمة t مباشرة من الجدول حيث

t [15, 0.05] = 1.753

ملاحظة: إذا كانت قيمة α غير موجودة في الجدول نستخدم العلاقة

$t\,[df, \alpha] = -t[df, 1-\alpha]$

مثال: احسب قيمة t فيما يلي:

1) t[29, 0.025]
2) t[29,0.001]
3) t[25,0.95]
4) t[3, 0.995]

الحل:

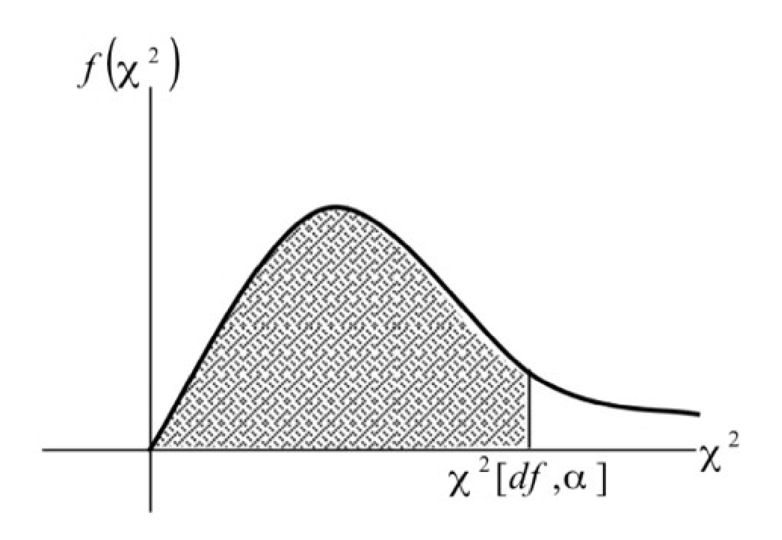
1) t[29,0.025] = 2.045
2) t[29, 0.001] =3.396
3) t[25,0.95] = -t [25,1-0.95]
= -t[25,0.05] = -1.708
4) t[3,0.995] = -t[3,0.005]
= -0.5841

3- توزيع كاي تربيع: χ^2 distribution

إن توزيع χ^2 من التوزيعات الاحتمالية المتصلة حيث يكون اقتران كثافته الاحتمالية معطى بالعلامة

$$f(\chi^2) = c\left(\chi^2\right)^{(\gamma-c)/2} \quad e^{-\frac{x^2}{2}} \qquad , \quad \chi^2 > 0$$

ومن هنا نرى أن منحنى χ^2 يكون في الموجب فقط حيث يكون منحناه على الصورة.



ويعطي الجدول قيمة χ^2 التي على يسارها مساحة α ودرجات حرية df.

مثال:

احسب قيمة χ^2 في كل مما يلي:

1) χ^2 [20,0.99]
2) χ^2[15.0.05]
3) χ^2[2,0.975]

الحل:

1) χ^2 [20,0.99] = 8.2604
2) χ^2 [15,0.05] = 24.9958
3) χ^2[2,0.975] = 0.0506356

تمارين

1- إذا كان الوسط الحسابي لعلامات شعبة ما في مادة الرياضيات (60) والانحراف المعياري (5)، وكان الوسط الحسابي لهذه الشعبة في مادة الفيزياء (70) والانحراف المعياري (10) فإذا كانت علامتي أحد طلاب هذه الشعبة في الرياضيات والفيزياء (65)، (75) على الترتيب فهل تحصيل الطالب في الفيزياء أفضل من الرياضيات؟

2- أوجد نسبة المساحة تحت المنحنى الطبيعي المعياري فيما يلي:

a) $P(Z<2.25)$　　b) $P(Z>1.7)$

c) $P(|Z|<1)$　　d) $P(|Z|>2)$

e) $P(-1.2<Z<1.78)$　　f) $P(Z<-1.8)$

3- أوجد نسبة المساحة تحت المنحنى الطبيعي فيما يلي:

a) $P(0.25 < Z <1.7)$　　b) $P(-1.5<Z<-0.05)$

c) $P(4Z<1.6)$　　d) $P(Z^2<Z)$

e) $P(Z^2<Z+2)$　　f) $P(Z^2>Z+2)$

g) $P(1<|Z|<2)$

4- أوجد قيمة a التي تحقق العلاقات في كل مما يلي:

a) $P(Z>a) = 0.0606$　　b) $P(Z<a)= 0.9938$

c) $P(|Z|<a)= 0.663$　　d) $P(Z<a)= 0.287$

5- إذا كانت $x_1, x_2, ..., x_n$ مجموعة من المشاهدات وسطها الحسابي $\bar{x}$، وانحرافها المعياري σ. وكانت $Z_1, Z_2, ..., Z_n$ العلامات المعيارية المناظرة لتلك المشاهدات. فأثبت أن:

a) $\overline{Z} = 0 \Rightarrow \sum_{r=1}^{n} Z_r = 0$

b) $\sigma_z = 1$□

إرشاد: $Z = \dfrac{X - \overline{X}}{\sigma} = \dfrac{1}{\sigma} X - \dfrac{\overline{X}}{\sigma}$

6- يتخذ الزمن اللازم لإنهاء (1000) طالب امتحانهم توزيعاً طبيعياً وسطه الحسابي (60) دقيقة وانحرافه المعياري (10) دقائق.

a) أوجد عدد الطلبة الذين ينهون الامتحان خلال أول (50) دقيقة من بدء الامتحان.

b) ما هي الفترة الزمنية اللازمة حتى يكون (800) طالباً قد أنهوا امتحانهم خلالها.

7- إذا كان بيع أحد المتاجر اليومي يتخذ توزيعاً طبيعياً وسطه الحسابي (300) دينار وانحرافه المعياري (20) دينار. احسب النسبة المئوية للأيام التي يزيد فيها بيع المتجر عن (320) دينار لليوم الواحد.

8- قرية تتكون من (10000) أسرة، يتخذ دخل الأسرة في هذه القرية توزيعا طبيعياً وسطه الحسابي (150) ديناراً وانحرافه المعياري (30) ديناراً أوجد:

a) النسبة المئوية للأسر التي يزيد دخلها عن (195) دينار في هذه القرية.

b) عدد الأسر التي تتراوح دخولهم بين (120) دينار، (180) دينار.

9- إذا كان زمن التشغيل لأحد أنواع البطاريات الجافة يتخذ توزيعا طبيعيا بوسط حسابي (150) ساعة تشغيل، وانحراف معياري (25) ساعة تشغيل. احسب:

a) النسبة المئوية للبطاريات التي تعمل أكثر من (178) ساعة.

b) إذا اعتبرت البطارية التي يقل زمن تشغيلها عن 59 ساعة تالفة، فما هي النسبة المئوية للتالف من هذه البطاريات؟

c) P_{60}

10- إذا كانت علامات (80) ألف طالب في الثانوية العامة تتخذ توزيعاً طبيعياً وسطه الحسابي (55) وانحرافه المعياري (10)، احسب:

a) عدد الطلبة الناجحون إذا كانت علامة النجاح (50).

b) إذا قبلت الجامعات الرسمية على المعدل التنافسي عشرة آلاف طالباً، فما هو أقل معدل قبل حسب التنافس؟

c) الرتبة المئينية للعلامة (72).

d) المدى الربيعي للعلامات.

11- إذا كان سعر التداول لسهم إحدى الشركات في سنة 2003 يتخذ توزيعاً طبيعياً بوسط حسابي (1.25)ديناراً وبانحراف معياري (15) قرشاً، احسب:

a- عدد الأيام في ذلك العام والتي قل فيها سعر التداول عن (140) قرشاً.

b- إذا علمت أن القيمة الاسمية للسهم ديناراً واحداً، فما هو عدد الأيام من ذلك العام والتي زاد فيها سعر السهم عن قيمته الاسمية؟

12- أسيل وهيا وسهير طالبات في الصف العاشر. فإذا كانت علامات الشهرين والعلامات المعيارية المناظرة معطاة في الجدول التالي:

السم الطالبة	العلامة الخام	العلامة المعيارية
أسيل	18	1
هيا	15	-0.5
سهير	x	1.5

احسب قيمة x.

13- إذا كانت أطوال طلاب الصف الأول أساسي وسطها الحسابي عشرة أمثال انحرافها المعياري. وكان طول الطالب محمد (130cm)، يقابله طولا معيارية قدره $\left(\frac{2}{3}\right)$. احسب العلامة المعيارية لطول الطالب صهيب والبالغ (150cm).

14- في يوم الشجرة تم زراعة (500) شتلة حرجية في حرم جامعة البلقاء التطبيقية، فإذا كان احتمال نجاح الشتلة الواحدة 75% احسب:

a- احتمال نجاح جميع الشتلات.

b- احتمال نجاح نصف الشتلات.

c- توقع عدد الشتلات الناجحة.

15- مستشفى للولادة فيه (30) سيدة في حالة ولادة، إذا فرضنا أن كل سيدة ستضع طفلاً واحداً فقط. احسب:

a- احتمال أن تنجب (20) سيدة أطفالاً ذكوراً.

b- احتمال أن تكون جميع المواليد إناثاً.

c- توقع عدد المواليد الإناث.

16- X متغير عشوائي لتجربة ذات الحدين، إذا كان احتمال نجاح التجربة في المرة الواحدة 0.15 فما عدد مرات إجراء التجربة إذا كان E(X)=45.

17- X متغير عشوائي لتوزيع ذات الحديثن، E(X)=3، وكان احتمال نجاح التجربة في المرة الواحدة (0.25n) حيث (n) عدد مرات إجراء التجربة. احسب قيمة n.

18- تقدم طالب لامتحان مستوى اللغة الإنجليزية يتكون من مائة فقرة من نوع الاختيار من متعدد متساوية العلامة، وعدد بدائل كل فقرة خمس بدائل منها واحدة صحيحة فقط. فإذا علمت أن الطالب أجاب جميع الفقرات عشوائياً. احسب:

a- احتمال أن يجيب على 25 فقرة على الأقل إجابة صحيحة.

b- احتمال نجاح الطالب في الامتحان إذا كانت علامة النجاح 70%.

c- توقع عدد الإجابات الصحيحة.

19- إحدى شركات الكمبيوتر، ترسل رسائل عبر شبكة الإنترنت، فإذا كان احتمال وجود خطأ في أية رسالة 2%، وأرسلت الشركة في أحد الأيام (300) رسالة، احسب:

a- احتمال عدم وجود خطأ في 200 رسالة.

b- توقع عدد الرسائل التي تحوي أخطاء.

20- شرطي مرور يقف على تقاطع طرق يومياً من الساعة العاشرة صباحاً حتى الثانية عشرة ظهراً. فإذا كان يلاحظ مرور (5) سيارات سياحية يومياً خلال هذه الفترة. احسب احتمال مرور ست سيارات سياحة في نفس الفترة في يوم ما.

21- احسب:

a) t [4, 0.005]

b) t [45, 0.1]

c) t [25 , 0.95]

22- احسب

a) χ^2 [7, 0.05]

b) χ^2 [42, 0.99]

c) χ^2 [1, 0.1]

7

الوحدة السابعة

التقدير واختبار الفرضيات

Estimation and Testing Hypothesis

التقدير واختبار الفرضية

Estimation and Testing Hypothesis

سنتناول في هذه الوحدة موضوعين هامين في الاستدلال الإحصائي (Statistical inference) وهما:

أولاً: التقدير الإحصائي (Statistical Estimation)

كثير من الأحيان نحتاج إلى معلمة إحصائية متعلقة بالمجتمع الإحصائي. ولكن لأسباب مختلفة لا يمكننا الحصول عليها مباشرة فنلجأ إلى تقديرها باستخدام عينة مأخوذة من المجتمع فمثلاً إذا أردنا معرفة معدل عمر تشغيل بطاريات جافة من نوع معين فلا يمكننا إيجاد الوسط الحسابي لعمر صلاحية هذه البطاريات، لذلك نأخذ عينة مناسبة من خلالها تقدر الوسط الحسابي للعمر التشغيلي لها. وقد نقدر هذا الوسط بقيمة معينة كأن نقول العمر التشغيلي (20) ساعة. أو قد نعطي فترة تقديرية لهذا الوسط فنقول أن العمر التشغيلي يقع ضمن الفترة [15, 25].

والآن وقبل أن نتعرف على كيفية التقدير سواءً بقيمة أو بفترة. ستتعرف على المصطلحات التالية:

- **المعلمة الإحصائية (Statistical Parameter)**

وفي موضوعنا ستكون المعلمة الإحصائية هي أحد مقاييس المجتمع مثل الوسط الحسابي للمجتمع (μ) وتباين المجتمع (σ^2) ونسبة المجتمع (P).

والتي تقابل مقاييس العينة وهي على الترتيب $\bar{x}$ ، S^2 ، $\hat{p}$

- **مستوى الثقة (Confedence level):**

وهي احتمال وقوع المعلمة الإحصائية ضمن فترة معينة تسمى فترة الثقة (Confidence interval)

.

• المجتمع الإحصائي (Statistical population)

هو موضوع الدراسة وسنعتبره في هذه الوحدة يأخذ توزيعاً طبيعياً.

أ- التقدير النقطي Point Estimation

وهنا نقدر معالم المجتمع بمقاييس العينة فيقدر الوسط الحسابي للمجتمع (μ) بـ الوسط الحسابي للعينة $(\overline{x})$ وتباين المجتمع (σ^2) بتباين العينة (S^2) ونسبة المجتمع (p) بنسبة العينة $(\hat{p})$

ولكن هذا التقدير يكتنفه بعض السلبيات حيث لا يكون دقيقاً باحتمال يمكن الاعتماد عليه.

مثال:

أخذت عينة من مجتمع إحصائي وسطه الحسابي (μ) وانحرافه المعياري (σ) حيث كان الوسط الحسابي للعينة هو $(\overline{x} = 5)$، وتباينها ($S^2=4$). قدّر وسط المجتمع وانحرافه المعياري.

الحل: يقدر وسط المجتمع بـ $\mu = \overline{x}$

أي أن $\mu = 5$

ويقدر انحرافه المعياري بـ $\sigma = s$

أي أن $\sigma = 2$

مثال:

أخذت عينة من طلبة جامعة البترا حجمها (200) طالباً وكان عدد الطلبة المغتربين فيها (60) طالباً. قدر نسبة الطلبة المغتربون في الجامعة.

الحل:

$$p = \hat{p}$$

$$p = \frac{60}{200} = 0.3$$

ب- التقدير بفترة Interval Estimation

1- للعينات الكبيرة (n ≥ 30)

- **تقدير الوسط الحسابي (μ)**

يقدر الوسط الحسابي بفترة الثقة

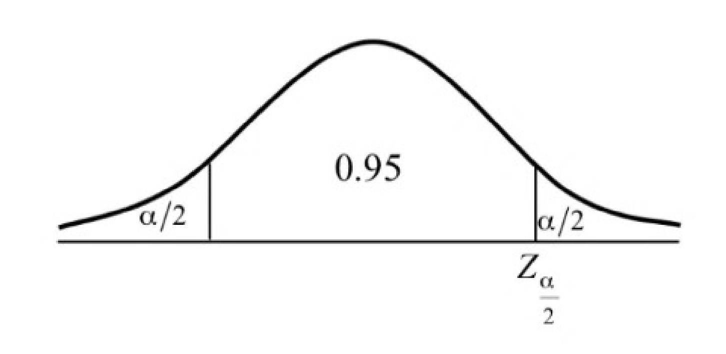
$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

ملاحظة: إذا كان تباين المجتمع غير معلوم تقدر σ بـ s.

مثال:

أخذت عينة حجمها (49) ووسطها (45) من مجتمع إحصائي يتخذ توزيعاً طبيعياً بانحراف معياري (3.5). جد فترة 95% ثقة لوسط المجتمع.

الحل:



نجد في البداية قيمة $Z_{\alpha/2}$.

حيث $\alpha = 1-0.95$

$= 0.05$

فتكون $\frac{\alpha}{2} = 0.025$

ونجد قيمة $Z_{\frac{\alpha}{2}}$ بالاستفادة من جدول

التوزيع الطبيعي المعياري حيث تكون القيمة المقابلة للمساحة 0.475

$$\therefore \quad Z_{\frac{\alpha}{2}} = 1.96$$

وبذلك تكون فترة الثقة عند مستوى دلالة 95% هي

$$45 - 1.96\frac{3.5}{\sqrt{49}} \le \mu \le 45 + 1.96\frac{3.5}{\sqrt{49}}$$

$$45 - 0.98 \le \mu \le 45 + 0.98$$

$$44.02 \le \mu \le 45.98$$

وهذا يعني أن احتمال وقوع وسط المجتمع بين القيمتين 45.98 , 44.02 هو 95%.

لتسهيل حل مسائل من هـذا النـوع سـنعطي قـيم $Z_{\alpha/2}$ لفـترات الثقـة عنـد مستويات دلالة شائعة ونلخصها في الجدول التالي:

Confidence Level	$Z_{\alpha/2}$
80%	1.28
90 %	1.645
95%	1.96
98%	2.33
99%	2.576

مثال:

إذا كان معدّل السحب اليـومي بالـدينار لوحـدة الـصراف الآلـي في أحـد فـروع البنك العربي مساوياً (385) دينار بانحراف معياري (27) دينار أوجد فـترة 99% ثقـة لمجتمع الأشخاص الذين يستخدمون وحدة الصراف الآلي في ذلـك الفـرع. إذا علمـت أن حجم العينة قيد الدراسة (900) عميل.

الحل:

بما أن حجم العينة كبيراً نـستخدم (S) بـدلاً مـن σ في القـانون. فيـصبح القـانون على الصورة

$$\bar{x} - Z_{\alpha/2}\frac{S}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{S}{\sqrt{n}}$$

فتكون الفترة المطلوبة هي

$$385 - (2.576)\frac{27}{\sqrt{900}} \le \mu \le 385 + (2.576)\frac{27}{\sqrt{900}}$$

$$385 - 2.318 \le \mu \le 385 + 2.318$$

$$382.682 \le \mu \le 387.318$$

- **تقدير الفرق بين وسطين $(\mu_1-\mu_2)$ (Estimation of mean difference)**

إذا كان لدينا مجتمعين مستقلين يتخذان توزيعاً طبيعياً وسطيهما الحسابيين μ_1 ، μ_2 وانحرافيهما σ_1^2 ، σ_2^2 على الترتيب. فإن فترة $(1-\alpha)\%$ ثقة للفرق بين وسطي المجتمعين $\mu_1-\mu_2$ هي

$$\left(\bar{x}_1-\bar{x}_2\right)-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\le \mu_1-\mu_2\le \left(\bar{x}_1-\bar{x}_2\right)+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

حيث n_1، n_2 حجم كل من العينتين على الترتيب

مثال:

أخذت عينتان من مشتركي شركتي اتصالات خلوية وأعطت المعلومات التالية:

تباين مجتمع الدقائق المستخدمة	معدل الدقائق المستخدمة	حجم العينة (n_i)
$\sigma_1^2=15$	$\bar{x}_1=75$	$n_1=100$
$\sigma_2^2=8$	$\bar{x}_2=80$	$n_2=81$

جد فترة 90% ثقة للفرق بين وسطين المجتمعين

الحل:

$$(75-80)-(1.645)\sqrt{\frac{15}{100}+\frac{8}{81}}\le \mu_1-\mu_2\le (75-80)+(1.645)\sqrt{\frac{15}{100}+\frac{8}{81}}$$

$$-5-0.822\quad \le \mu_1-\mu_2\le\quad -5+0.823$$

$$-5.823\quad \le \mu_1-\mu_2\le\quad -4.177$$

ملاحظة: إذا كانت σ_1 ، σ_2 غير معلومتين يمكن الاستعاضة عنهما بـ s_1، s_2.

- **فترة الثقة لنسبة المجتمع (p)**

(Interval Estimation of Population Proportion)

$$\hat{p} - Z_{a/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} \le p \le \hat{p} + Z_{a/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$$

ملاحظة:

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

مثال:

أخذت عينة حجمها (500) طالبة من مجتمع طالبـات الجامعـات الأردنيـة فكـان عدد المحجبات منهن (280) طالبة. أوجد فترة 95% ثقة لنسبة الطالبات المحجبـات في الجامعات الأردنية.

الحل:

$$n = 500$$

$$\hat{p} = \frac{280}{500} = 0.56$$

فتكون الفترة المطلوبة هي:

$$0.56 - 1.69\sqrt{\frac{(0.56)(0.44)}{500}} \le p \le 0.56 + 1.69\sqrt{\frac{(0.56)(0.44)}{500}}$$

$$0.56 - 0.044 \le p \le 0.56 + 0.044$$

$$0.516 \le p \le 0.604$$

- **تقدير فترة الثقة للفرق بين نسبتين**

(Interval estimation for proportion difference)

لمجتمعين مستقلين تكون فترة $(1-\alpha)\%$ ثقة للفرق بين نسبتين المجتمعين هي:

$$(\hat{p}_1 - \hat{p}_2) - Z_{a/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \le p_1 - p_2 \le (\hat{p}_1 - \hat{p}_2) - Z_{a/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

مثال:

مصنعين لأجهزة التلفزيون، أخذت عينة من إنتاج المصنع الأول حجمها (40) جهازاً وكانت نسبة المعيب فيها (2%) وأخذت عينة أخرى من إنتاج المصنع الثاني حجمها (60) جهازاً فكانت نسبة المعيب فيها 1.5% .

أوجد فترة 98% ثقة للفرق بين نسبتي الأجهزة الصالحة بين المصنعين.

الحل:

$$\hat{p}_1 = 0.98 \qquad \hat{p}_2 = 0.985$$

$$n_1 = 40 \qquad n_2 = 60$$

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{(0.98)(0.02)}{40} + \frac{(0.985)(0.015)}{60}}$$

$$= 0.027$$

فتكون فترة الثقة المطلوبة هي:

$$(0.98-0.985) - (2.33)\ (0.027) \le p_1-p_2 \le (0.98-0.985) + (2.33)\ (0.027)$$

$$-0.005 - 0.063 \le p_1-p_2 \le -0.005 + 0.063$$

$$-0.068 \le p_1-p_2 \le 0.058$$

2- للعينات الصغيرة (n<30)

- **فترة الثقة للوسط الحسابي**

$$\bar{x} - t_{a/2}\frac{S}{\sqrt{n}} \le m \le \bar{x} - t_{a/2}\frac{S}{\sqrt{n}}$$

> ملاحظة: إذا كانت σ معلومة نستخدم توزيع (Z) بدلاً من توزيع t (مثل حالة العينات الكبيرة)

مثال:

أخذت عينة حجمها (9) عبوات من إنتاج إحدى آلات تعبئة العصير للمصنع ما فكان حجم العصير في كل منها بالمليمتر 245، 248، 251، 247، 253، 243، 249، 253، 252.

أوجد فترة 95% ثقة لمعدل حجم إنتاج الآلة.

الحل:

$$\bar{x} = \frac{\sum x}{n} = 249$$

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

$$s = \sqrt{12.75} = 3.57$$

$$t_{\frac{\alpha}{2}} = t\left[n - 1 , \frac{\alpha}{2}\right]$$

$$= t\,[8 , 0.025]$$

$$= 2.306$$

فتكون فترة الثقة المطلوبة:

$$249 - 2.306\frac{3.57}{\sqrt{9}} \leq \mu \leq 249 + 2.306\frac{3.57}{\sqrt{9}}\,249$$

$$249 - 2.74 \leq \mu \leq 249 + 2.74$$

$$246.26 \leq \mu \leq 251.74$$

- **فترات الثقة للفرق بين وسطين**

لمجتمعين مستقلين تكون فترة الثقة للفرق بين وسطني بمستوى ثقة $(1\text{-}\alpha)\%$ هي

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{a}{2}}\, S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq m_1 - m_2 \leq (\bar{x}_1 - \bar{x}_2) + t_{\frac{a}{2}}\, S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

حيث

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$t_{\frac{a}{2}} = t\left[n_1 + n_2 - 2 , \frac{a}{2}\right]$$

مثال:

أخذت عينتان من مجتمعين متجانس تباينهما وأعطت النتائج التالية:

	العينة الأول	العينة الثانية
حجم العينة	$n_1 = 17$	$n_2=25$
الوسط الحسابي	$\bar{x}_1 = 120$	$\bar{x}_2 = 112$
الانحراف المعياري	$s_1 = \sqrt{10}$	$s_2 = 5$

أوجد فترة 90% ثقة للفرق بين وسطي المجتمعين.

الحل:

نحسب

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(16)(10) + (24)(25)}{40}$$

$$S^2 = 19 \Rightarrow S = 4.36$$

$$t_{\frac{\alpha}{2}} = t\left[n_1 + n_2 - 2\ , \frac{\alpha}{2}\right]$$

$$= t\ [40\ ,\ 0.05] = 1.684$$

$$S\ \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = (4.36)(0.314) = 1.37$$

فتكون فترة الثقة المطلوبة

$$(120-112) - (1.684)\ (1.37) \le \mu_1 - \mu_2 \le (120-112) + (1.684)\ (1.37)$$

$$8 - 2.31 \le \mu_1 - \mu_2 \le 8 + 2.31$$

$$5.69 \le \mu_1 - \mu_2 \le 10.31$$

ثانياً: اختبار الفرضيات (Testing Hypotheses)

يعتبر موضوع اختبار الفرضيات من المواضيع الهامة في الإحصاء الاستدلالي حيث يحتاجه كل باحث مهما كان تخصصه وفي العادة تستخدم فرضيتان الأولى تسمى الفرضية الصفرية (العدم) (null hypothesis) وعادة يرمز لها بالرمز (H_0) والأخرى تسمى الفرضية البديلة (alternative hypothesis) ويرمز لها بالرمز (H_1). ويكون القرار الإحصائي بقبول أو رفض الفرضية الصفرية.

والفرضية الصفرية تكون بالوضع المحايد.

أنواع الخطأ

1- الخطأ من النوع الأولى (type I error)

وهو رفض الفرضية الصفرية عندما تكون صحيحة.

2- الخطأ من النوع الثاني (Type II error)

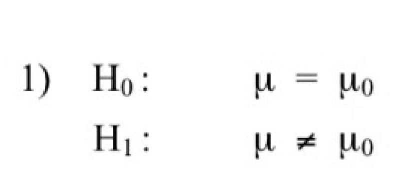
وهو قبول الفرضية الصفرية وهي خاطئة.

هنالك عدة اختبارات منها ما هو متعلق بالوسط أو الفرق بين وسطين، النسبة والفرق بين نسبتين. وسندرس هذه الاختبارات بشيء من التفصيل تالياً.

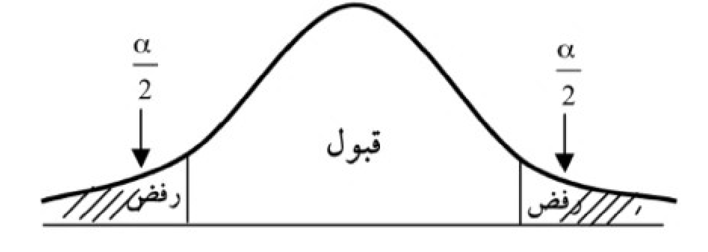
- **اختبار الفرضيات المتعلق بالوسط الحسابي**

ويكون الهدف منه اختبار فيما إذا كان الوسط الحسابي يساوي أو أكبر أو أصغر من قيمة ما.

وتكون الفرضيات هذه الحالات كما يلي وعلى الترتيب

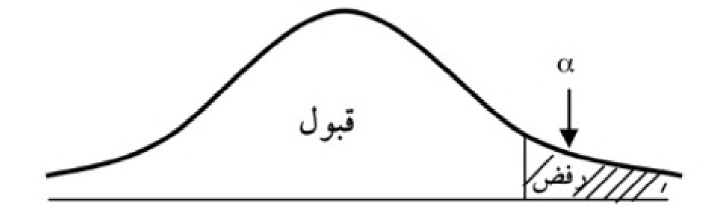

1) $H_0: \quad \mu = \mu_0$

$H_1: \quad \mu \neq \mu_0$



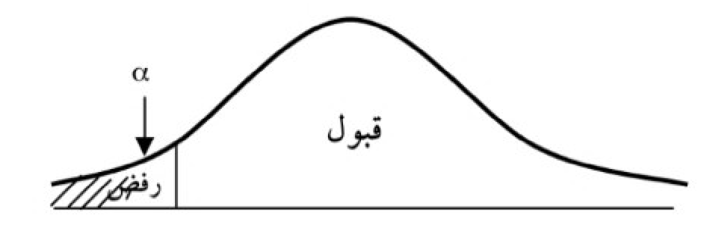
2) $H_0: \quad \mu = \mu_0$

$H_1: \quad \mu > \mu_0$



3) $H_0: \quad \mu = \mu_0$

$H_1: \quad \mu < \mu_0$



حيث تمثل α: احتمال رفض الفرضية الصفرية وتسمى **مستوى الدلالة**.

أ- اختبار الوسط للعينات الكبيرة

لإجراء هذا الاختبار بمستوى دلالة α

1. نكتب الفرضيات الإحصائية المناسبة.
2. نجد قيمة الإحصائية Z بالقانون.

$$Z = \frac{\bar{x} - m_0}{s/\sqrt{n}}$$

3. نحدد منطقة الرفض والقبول باستخدام قيمة Z الحرجة (Critical point). وتسمى أيضا Z الجدولية.
4. وإذا وقع الإحصائي Z في منطقة القبول نقبل الفرضية الصفرية ونرفض البديلة. وغير ذلك نرفض الفرضية الصفرية ونقبل البديلة.

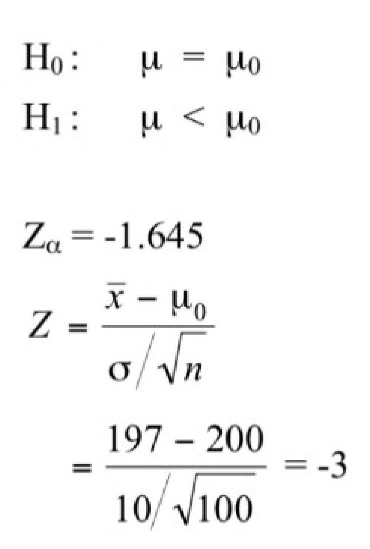
مثال:

إذا حددت مديرية المواصفات والمقاييس وزن رغيف الخبز بـ (200gm). بانحراف معياري (10gm) فإذا أخذ أحد المفتشين عينة من مخبز معين مكونة من (100) رغيف فكان الوسط الحسابي لأوزانها (197gm).

فهـل هـذا المخبـز يعتـبر مخـالف للمواصـفات والمقـاييس عنـد مـستوى دلالـة (α=0.05)؟

الحل:

سنعتبر أن الشخص مخالف إذا كان وزن الرغيف أقل من (200gm).



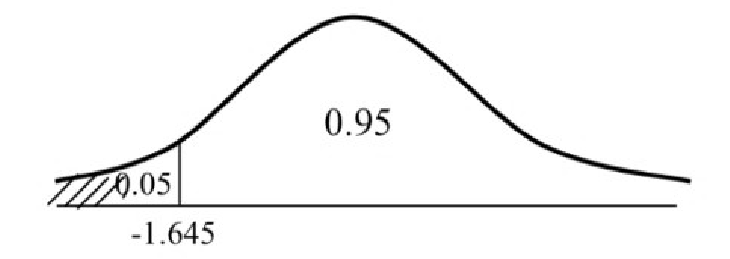
$H_0: \quad \mu = \mu_0$

$H_1: \quad \mu < \mu_0$

$Z_\alpha = -1.645$

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

$$= \frac{197 - 200}{10/\sqrt{100}} = -3$$



وتقع قيمة الإحصائي ضمن منطقة الرفض لذلك نرفض الفرضية الـصفرية ونقبـل البديلة أي أن وزن الرغيف يقل عن الوزن المقـرر مـن مديريـة المواصـفات والمقـاييس عند مستوى الدلالة المعطى لذلك فهو مخالف.

مثال:

ينـتج مـصنع أقـراص مرنـة (floppy disk) للحاسـوب ذات القطـر (3.500 inch) بانحراف معياري (0.02inch). أخذت عينة من إنتاج المـصنع مكونـة مـن (64) قرصـاً وقيست أقطارها فأعطت متوسط حسابي مقداره 3.505. فهل تعتبر هذه العينة ملائمـة لهذا النوع من الأقراص عند مستوى دلالة (α=0.01).

الحل:

$H_0: \quad \mu = 3.500$

$H_1: \quad \mu \neq 3.500$

$$Z = \frac{3.505 - 3.500}{0.02/\sqrt{64}}$$

$$Z = 2$$

0.99
0.005
0.005
-2.576
2.576

تقع هذه القيمة في منطقة القبول. فلذلك نقبل الفرضية الصفرية ونرفض البديلة. أي أن الأقراص ملائمة عند مستوى الدلالة المعطى.

ملاحظة: إذا كانت σ غير معلومة نستخدم (S) كقيمة تقديرية حيث تصبح قيمة الإحصائي:

$$Z = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

ب- اختبار الوسط في حالة العينات الصغيرة وتباين المجتمع غير معلوم:

نستخدم هنا توزيع t بدرجات حرية (n-1) بدلاً من Z. ويكون الإحصائي

$$t = \frac{\overline{x} - m_0}{S/\sqrt{n}}$$

مثال:

إذا كان معدل نسبة النيكوتين المثبتة على أحد أنواع السجائر هي (0.7 mlg). أخذ (16) سيجارة من هذا النوع فكان الوسط الحسابي لنسبة النيكوتين تساوي (0.75mlg) بانحراف معياري (0.04 mlg). فهل تعتبر هذه السجائر ذات معدل نسبة نيكوتين أعلى من المثبت على علبة السجائر عند مستوى دلالة (α=0.05).

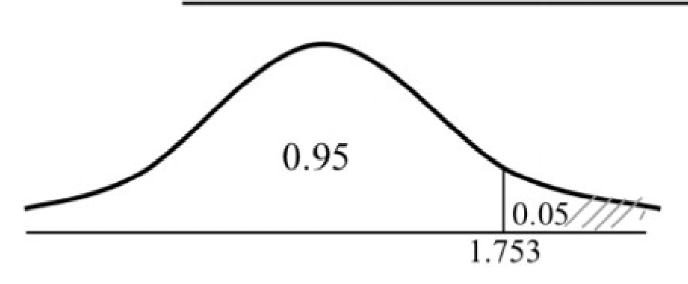
الحل:

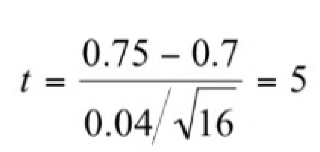
$H_0: \quad \mu = 0.7$

$H_1: \quad \mu > 0.7$

$t_\alpha = t\,[15\ ,\ 0.5]$

$= 1.753$





$$t = \frac{0.75 - 0.7}{0.04 \big/ \sqrt{16}} = 5$$

وتقع في منطقة الرفض أي أن معدل نسبة النيكوتين أعلى من المثبت على علبة الدخان.

- **اختبار الفرضيات المتعلقة بالفرق بين وسطين**

أ- العينات الكبيرة:

إذا كان لدينا مجتمعين مستقلين تباين كل منهما على الترتيب σ_1, σ_2 بوسطين حسابيين μ_1, μ_2 فيمكن أن نختبر إحدى الحالات التالية:

1) $H_0: \quad \mu_1 = \mu_2 \quad \Leftrightarrow \quad \mu_1 - \mu_2 = 0$
   $H_1: \quad \mu_1 \neq \mu_2 \quad \Leftrightarrow \quad \mu_1 - \mu_2 \neq 0$

2) $H_0: \quad \mu_1 = \mu_2 \quad \Leftrightarrow \quad \mu_1 - \mu_2 = 0$
   $H_1: \quad \mu_1 > \mu_2 \quad \Leftrightarrow \quad \mu_1 - \mu_2 > 0$

3) $H_0: \quad \mu_1 = \mu_2 \quad \Leftrightarrow \quad \mu_1 - \mu_2 = 0$
   $H_1: \quad \mu_1 < \mu_2 \quad \Leftrightarrow \quad \mu_1 - \mu_2 < 0$

ويكون الإحصائي في جميع الحالات

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (m_1 - m_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\text{or} \quad Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

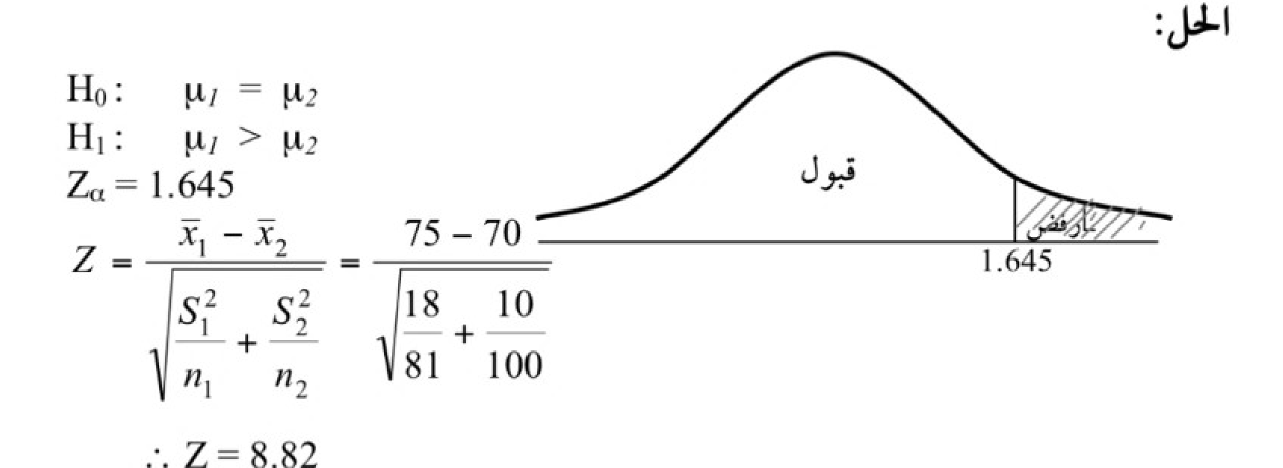
مثال:

يدعي أحد الباحثين أن متوسط علامات طلبة الجامعات الحكومية في مادة الإحصاء أفضل من طلبة الجامعات الخاصة في نفس المادة. فإذا أخذت عينتان من الجامعات الحكومية والخاصة وأعطت النتائج التالية:

	عينة الجامعات الحكومية	عينة الجامعات الخاصة
حجم العينة	$n_1 = 81$	$n_2 = 100$
الوسط الحسابي	$\bar{x}_1 = 75$	$\bar{x}_2 = 70$
التباين	$s_1^2 = 18$	$s_2^2 = 10$

اختبر ادعاء الباحث على مستوى دلالة $(\alpha=0.05)$.

الحل:



$H_0: \quad \mu_1 = \mu_2$

$H_1: \quad \mu_1 > \mu_2$

$Z_\alpha = 1.645$

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{75 - 70}{\sqrt{\frac{18}{81} + \frac{10}{100}}}$$

$\therefore Z = 8.82$

وتقع هذه القيمة في منطقة الرفض أي أن ادعاء الباحث صحيح.

> **ملاحظة:** في المثال السابق استخدم S_1^2، S_2^2 بدلاً من σ_1^2، σ_2^2 على الترتيب مع استخدام الإحصائي Z وذلك لأن العينتان كبيرتان.

ب- للعينات الصغيرة والتباين غير معلوم

إذا أخذت عينتان صغيرتان (اقل من 30) مستقلتان من مجتمعين وسطيهما على الترتيب μ_1, μ_2، وتباينهما غير معلومين فإننا نستخدم توزيع t بدلاً من Z بدرجات حرية (n_1+n_2-2) ويكون الإحصائي هو

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

حيث

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

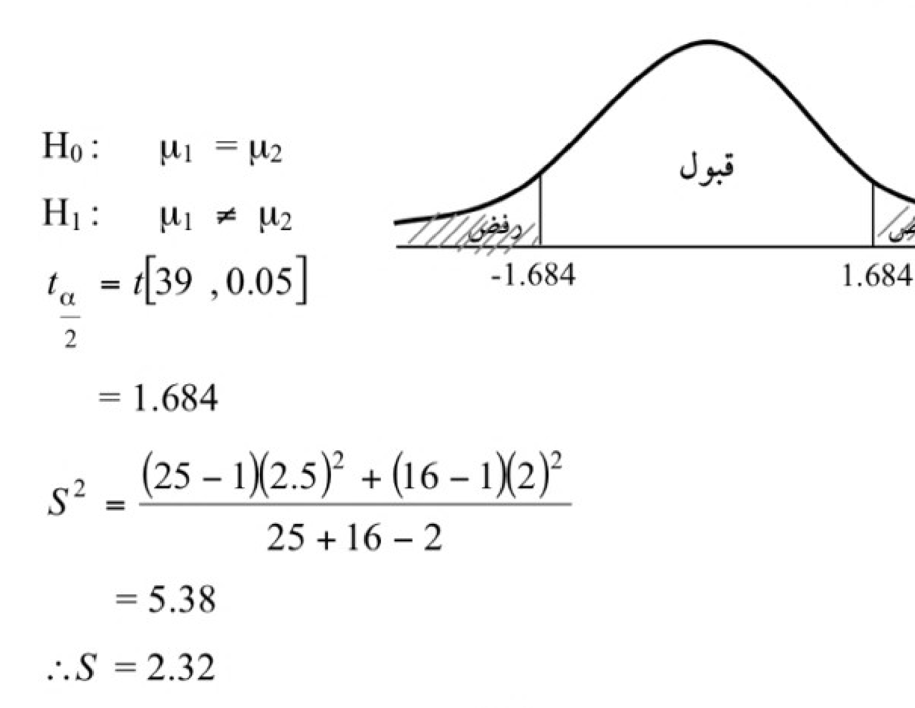
مثال:

أخذت عينتان من إنتاج مصنعين لأجهزة التلفزيون وأعطت النتائج التالية:

	عينة المصنع الأول	عينة المصنع الثاني
حجم العينة	$n_1 = 25$	$n_2 = 16$
متوسط العمر التشغيلي بالسنة	$\bar{x}_1 = 17$	$\bar{x}_2 = 15$
الانحراف المعياري	$s_1 = 2.5$	$s_2 = 2$

بناء على العينتين هل يوجد فرق بين متوسطي العمر التشغيلي لإنتاج المصنعين على مستوى دلالة (α=0.1).

الحل:



$H_0: \quad \mu_1 = \mu_2$

$H_1: \quad \mu_1 \neq \mu_2$

$$t_{\frac{\alpha}{2}} = t[39\ ,0.05]$$

$$= 1.684$$

$$S^2 = \frac{(25-1)(2.5)^2 + (16-1)(2)^2}{25+16-2}$$

$$= 5.38$$

$$\therefore S = 2.32$$

$$\therefore t = \frac{17-15}{2.32\sqrt{\frac{1}{25}+\frac{1}{16}}}$$

$$= \frac{2}{(2.32)(0.32)}$$

$$t = 2.7$$

وتقع في منطقة الرفض

إذن يوجد فرق ذو دلالة بين متوسطي العمر التشغيلي لإنتاج المصنعين.

ملاحظة: إذا كان تباينا المجتمعين معلومين نستخدم توزيع Z بدلاً من t.

- **اختبار الفرضيات المتعلق بالنسبة**

في هذه الحالة تكون الفرضيات الممكنة

1) $H_0: \quad P = P_0$
   $H_1: \quad P \neq P_0$

2) $H_0: \quad P = P_0$
   $H_1: \quad P > P_0$

3) $H_0: \quad P = P_0$
   $H_1: \quad P < P_0$

الإحصائي يكون

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

مثال:

مدير يعاقب سكرتيرته إذا كانت نسبة الخطأ في الكتب التي تطبعها أكثر من 5% فإذا أخذت عينة مكونة من 50 كتاب من طباعة السكرتيرة ووجد أن ثلاث كتب منها تحوي أخطاء.

فهل هذا يعني أن السكرتيرة تستحق العقاب عند مستوى دلالة (α=0.05).

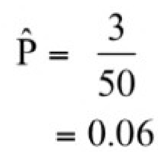
الحل:

قبول

رفض

1.645

$$H_0: \quad p = 0.05$$
$$H_1: \quad p > 0.05$$
$$Z_\alpha = 1.645$$



$$\hat{P} = \frac{3}{50} = 0.06$$

$$Z = \frac{0.06 - 0.05}{\sqrt{\frac{(0.05)(0.95)}{50}}} = \frac{0.01}{0.031} = 0.32$$

هذه القيمة تقع ضمن منطقة القبول وبالتالي لا تستحق السكرتيرة العقاب.

- **اختبار الفرضيات المتعلقة بالفرق بين نسبتين**

الفرضيات الممكنة:

1) $H_0: \quad P_1 = P_2$
$H_1: \quad P_1 \neq P_2$

2) $H_0: \quad P_1 = P_2$
$H_1: \quad P_1 > P_2$

3) $H_0: \quad P_1 = P_2$
$H_1: \quad P_1 < P_2$

والإحصائي يكون

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

حيث
$$\hat{p} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2}$$

مثال:

أخذت عينتان من الرجال والنساء في مجتمع ما فأعطت النتائج التالية:

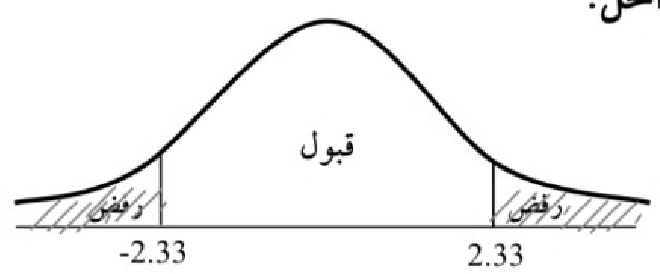
	عينة الرجال	عينة النساء
حجم العينة	$n_1 = 200$	$n_2 = 200$
عدد الذين يستخدمون الهاتف الخلوي	160	210

فهل هذه النتائج تدل على أن هنالك اختلاف بين نسبة الرجال الذين يستخدمون الخلوي عن نسبة النساء.

اختبر ذلك على مستوى دلالة ($\alpha=0.02$)



الحل:



$H_0: \quad P_1 = P_2$

$H_1: \quad P_1 \neq P_2$

$$Z_{\frac{\alpha}{2}} = 2.33$$

$$\hat{p}_1 = \frac{160}{200} = 0.8$$

$$\hat{p}_1 = \frac{210}{300} = 0.7$$

$$\hat{p}_2 = \frac{160 + 210}{500} = 0.74$$

$$Z = \frac{0.8 - 0.7}{\sqrt{(0.74)(0.36)\left(\frac{1}{200} + \frac{1}{300}\right)}}$$

$$= \frac{0.1}{0.047} = 2.13$$

وتقع في منطقة القبول أي أنه لا يوجد فرق بدلالة إحصائية بين النسبتين.

تمارين

1- أخذت عينة حجمها (64) من مجتمع إحصائي وسطه الحسابي (μ) وانحرافه المعياري (10)،إذا كان الوسط الحسابي للعينة يساوي 70. أوجد

1. تقدير نقطي ثقة لوسط المجتمع.
2. فترة 80% ثقة لوسط المجتمع

2- أخذ عينة حجمها (25) من مجتمع إحصائي وسطه الحسابي μ من مجتمع إحصائي وسطه الحسابي μ وانحرافه المعياري σ. فإذا كان الوسط الحسابي للعينة (17) وتباينها (4). أوجد:

1. فترة 90% ثقة لوسط المجتمع μ.
2. فترة 98% ثقة لوسط المجتمع μ.

3- أخذت عينتان من مجتمعين مستقلين وأعطت النتائج التالية:

	العينة الأولى	العينة الثانية
حجم العينة	$n_1 = 64$	$n_2 = 36$
الوسط الحسابي	$\bar{x}_1 = 50$	$\bar{x}_2 = 18$
الانحراف المعياري	$s_1 = 4$	$s_2 = 5$

جد فترة 95% ثقة للفرق بين الوسطين

a-(μ_1- μ_2)   b -(μ_2 -μ_1)

4- أخذت عينتان من مجتمعين مستقلين وأعطت النتائج التالية

	العينة الأولى	العينة الثانية
حجم العينة	$n_1 = 15$	$n_2 = 24$
الوسط الحسابي	$\bar{x}_1 = 123$	$\bar{x}_2 = 127$
التباين	$s_1 = 38$	$s_2 = 40$

جد فترة 99% ثقة للفرق بين الوسطين $(\mu_1 - \mu_2)$

5- إذا كان معدل دخل (10) أسر أخذت عشوائياً من مدينة عمان هو (200) دينار. وكان معدل دخل (15) أسرة أخذت عشوائياً من مدينة الزرقاء هو (175) دينار بانحراف معياري (10) دنانير. جد فترة 95% ثقة للفرق بين متوسطي الدخل لأسر مدينة عمان والزرقاء.

6- أخذت عينة من طلاب الجامعة الأردنية حجمها (400) طالباً فوجد أن (250) طالباً منهم يمتلكون سيارات خاصة. اكتب فترة 90% ثقة لنسبة الطلبة الذين يملكون سيارات خاصة في الجامعة.

7- أخذت عينة من السياح الذين يزورون مدينة العقبة الأردنية حجمها (500) سائحاً فكانت نسبة الأجانب منهم (25%) وأخذت عينة أخرى من السياح الذين يزورون مدينة شرم الشيخ المصرية حجمها (500) سائح وكان نسبة الأجانب منهم (40%). جد فترة 99% ثقة للفرق بين النسبتين.

8- أخذت عينة حجمها (225) من مجتمع إحصائي يتخذ توزيعاً طبيعياً بوسط حسابي μ وانحراف معياري (8). وكان الوسط الحسابي للعينة (35). اختبر الفرضية: $H_0: \mu=34$ مقابل $H_1: \mu \neq 34$ على مستوى دلالة $(\alpha=0.02)$

9- أخذت عينتان عشوائيتان من نوعين من المصابيح الأولى حجمها (40) مصباحاً بمتوسط ساعات تشغيلية مقداره (5000) ساعة بانحراف معياري (250) ساعة والثانية حجمها (50) مصباحاً بمتوسط ساعات تشغيلية مقداره (4500) بانحراف معياري (200) ساعة فهل يوجد فرق ذو دلالة إحصائية بين متوسطي الساعات التشغيلية على مستوى دلالة (0.05).

10- إذا كانت أطوال طلاب الصف الأول الأساسي تتخذ توزيعاً طبيعياً بوسط حسابي μ وانحراف معياري σ. أخذت عينة عشوائياً من طلاب الصف الأول الأساسي فكانت أطوالهم 110، 127، 119، 118، 125، 120، 101، 109. اختبر الفرضية القائلة بأن متوسط أطوال الصف الأول الأساسي أقل من (120 cm) على مستوى دلالة (α=0.01).

11- إذا كانت نسبة الشفاء من مرض معين إذا استخدم العلاج (A) هو (93٪). ثم أنتج أحد مصانع الأدوية نوعاً آخر من العلاج (B). وادعى أن هذا العلاج له نسبة شفاء أكبر من النوع الأول. فأخذت عينة مكونة من (120) مريض طبق عليهم العلاج B فشفي منهم (114) مريض. فهل ادعاء المصنع صحيح على مستوى دلالة (α=0.01).

12- أخذت عينتان من الذكور، الإناث حجماهما على الترتيب 80، 75 فكان عدد المصابين بالسرطان من عينة الذكور (4) وعدد المصابات بالسرطان من عينة الإناث (3). اختبر الادعاء القائل بأن نسبة المصابين بالسرطان من الذكور أعلى منها من الإناث عند مستوى دلالة (α=0.05).



الوحدة الثامنة

الأرقام القياسية

Index Numbers

الأرقام القياسية
Index Numbers

مفهوم الرقم القياسي

تعريف:

الرقم القياسي هو أداة لقياس التغير النسبي أو النسبي المئوي في قيم الظواهر من زمن إلى آخر أو من مكان إلى آخر. ويسمى الزمن أو المكان الأول بالأساس ويسمى الزمن أو المكان الثاني بالمقارن.

مثال:

إذا كان سعر كيلو الخبز سنة 1995 (15) قرش وفي سنة 2000 عشرين قرش. فما هو الرقم القياسي لسعر الخبز سنة 2000 باعتبار سنة 1995 الأساس؟

الحل:

$$\text{الرقم القياسي} = \frac{\text{سعر سنة المقارنة}}{\text{سعر سنة الأساس}} \times 100\%$$

$$= \frac{\text{سعر سنة 2000}}{\text{سعر سنة 1995}} \times 100\%$$

$$I = \frac{20}{15} \times 100\%$$

$$= 133.3\%$$

وهذا الرقم يعني أن كمية الخبز التي كان ثمنها عام 1995 مئة قرش أصبح ثمنها في عام 2000 تقريباً (133) قرشاً.

فوائد الرقم القياسي

1- معرفة نسبة التغير في ظاهرة ما من مكان لآخر أو من زمن لآخر.
2- معرفة الدخل الحقيقي للفرد أو ما يسمى بالقوة الشرائية لدخل الفرد.
3- معرفة نسبة الزيادة في الإنتاج من زمن لآخر.

مثال:

إذا كان الرقم القياسي لدخل الفرد عام 2002 باعتبار سنة 2000 الأساس هو 1.5، والرقم القياسي لتكاليف المعيشة في عام 2002 باعتبار سنة 2000 الأساس هو 3، فما هي القوة الشرائية لدخل الفرد في عام 2002 باعتبار عام 2000 هي الأساس.

الحل:

$$\text{القوة الشرائية لدخل الفرد} = \frac{\text{الرقم القياسي لدخل الفرد}}{\text{الرقم القياسي لتكاليف المعيشة}} \times 100\%$$

$$= \frac{1.5}{3} \times 100\% = 50\%$$

أي أن دخل الفرد قد نقص بنسبة 50٪ ما بين عام 2000 وعام 2002.

الرقم القياسي البسيط: Simple Index Number

وهناك نوعين من الأرقام القياسية البسيطة

1- الرقم القياسي البسيط للأسعار:

a- الرقم القياسي التجميعي البسيط للأسعار = $\frac{\text{مجموع أسعار سنة المقارنة}}{\text{مجموع أسعار سنة الأساس}} \times 100\%$

$$I_c(P) = \frac{\sum P_1}{\sum P_0} \times 100\%$$

حيث: سعر سنة المقارنة $= P_1$

سعر سنة الأساس $= P_0$

b- الرقم القياسي النسبي للأسعار = $\frac{1}{r}$ × مجموع $\left(\frac{\text{أسعار سنة المقارنة}}{\text{أسعار سنة الأساس}}\right)$×100٪

عدد الظواهر =r ,
$$I_p(P)=\frac{1}{r}\left(\sum\frac{P_1}{P_0}\right)\times 100\%$$

ملاحظة: الرقم القياسي النسبي هو الوسط الحسابي للأرقام القياسية للظواهر الداخلة في حسابه.

مثال:

إذا كان سعر بيع الوحدة لإنتاج ثلاثة مصانع في عامي 2004، 2003 معطاة في الجدول التالي:

المصنع	السعر (دينار للوحدة)	
	2003	**2004**
1	35	36
2	30	32
3	31	33

احسب الرقم القياسي التجميعي البسيط، والرقم القياسي النسبي البسيط للأسعار عام 2004 باعتبار عام 2002 هي الأساس.

الحل:

$$I_c(P)=\frac{\sum P_1}{\sum P_0}\times 100\%$$

$$=\frac{36+32+33}{35+30+31}\times 100\%$$

$$I_c(P)=\frac{101}{96}\times 100\%=105.21\%$$

$$I_p(P)=\frac{1}{r}\left(\sum\frac{P_1}{P_0}\right)\times 100\%$$

$$= \frac{1}{3}\left(\frac{36}{35} + \frac{32}{30} + \frac{33}{31}\right) \times 100\%$$

$$= \frac{1}{3}(1.03 + 1.07 + 1.06) \times 100\%$$

$I_P(P)$= 105,31%

2- الرقم القياسي البسيط للكميات: وهو رقمين:

a- الرقم القياسي التجميعي البسيط للكميات= $\frac{\text{مجموع كميات سنة المقارنة}}{\text{مجموع كميات سنة الأساس}} \times 100\%$

$$I_c(Q) = \frac{\sum Q_1}{\sum Q_0} \times 100\%$$

حيث:

كمية سنة المقارنة = Q_1

كمية سنة الأساس= Q_0

b- الرقم القياسي النسبي للكميات = $\frac{1}{r}\times$ مجموع $\left(\frac{\text{كمية سنة المقارنة}}{\text{كمية سنة الأساس}}\right) \times 100\%$

$$I_P(Q) = \frac{1}{r}\left(\sum \frac{Q_1}{Q_0}\right) \times 100\%$$

عدد الظواهر = r

مثال:

في متجر لبيع المواد الاستهلاكية إذا كانت كميات المواد المباعة بالطن في عامي 2003 , 2000 هي كما يلي:

المادة	كمية عام 2000	كمية عام 2003
سكر	500	560
رز	450	480
طحين	650	580
حليب	700	730

أوجد الرقم القياسي التجميعي والنسبي البسيط لكميات عام 2003 باعتبار عـام 2000 هو الأساس.

الحل:

$$I_c(Q) = \frac{\sum Q_1}{\sum Q_0} \times 100\%$$
$$= \frac{560 + 480 + 580 + 730}{500 + 450 + 650 + 700}$$
$$= \frac{2350}{2300} \times 100\%$$
$$= 102.17\%$$

$$I_P(Q) = \frac{1}{r}\left(\sum \frac{Q_1}{Q_0}\right) \times 100\%$$
$$= \frac{1}{4}\left(\frac{560}{500} + \frac{480}{450} + \frac{580}{650} + \frac{730}{700}\right) \times 100\%$$
$$= \frac{1}{4}(1.12 + 1.07 + 0.89 + 1.04) \times 100\%$$
$$= \frac{1}{4}(4.12)100\% = 103\%$$

الأرقام القياسية المرجحة (Weighted Index Number)

وهناك نوعان من الأرقام القياسية المرجحة

1- الرقم القياسي المرجح للأسعار: وهو ثلاثة أرقام:

a) الـرقم القياسـي للأسـعار والمـرجح بكميـات سـنة الأسـاس (رقـم لاسـبير للأسعار).

$$\text{رقم لاسبير للأسعار} = \frac{\text{مجموع (أسعار سنة المقارنة} \times \text{كميات سنة الأساس)}}{\text{مجموع (أسعار سنة الأساس} \times \text{كميات سنة الأساس)}} \times 100\%$$

$$laspeyre(P) = \frac{\sum P_1Q_0}{P_0Q_0} \times 100\%$$

b) الرقم القياسي للأسعار والمرجح بكميات سنة المقارنة (رقم باش للأسعار).

رقم باش للأسعار= $\frac{\text{مجموع (أسعار سنة المقارنة×كميات سنة المقارنة)}}{\text{مجموع (أسعار سنة الأساس×كميات سنة المقارنة)}}$ × 100٪

$$paasche(P) = \frac{\sum P_1Q_1}{P_0Q_1} \times 100\%$$

c) الرقم القياسي الأمثل للأسعار: (رقم فشر للأسعار).

$$Fisher(p) = \sqrt{Laspeyre(P) \times Paasche(P)}\ \%$$

أي هو الوسط الهندسي لرقمي لاسبير وباش للأسعار.

مثال:

الجدول التالي يمثل أسعار وكميات مبيعات أربع سلع التي بيعت عامي 2002، 2004.

السلعة	سعر الوحدة		كمية المبيعات	
	عام 2002	عام 2004	عام 2002	عام 2004
a	28	40	200	250
b	16	20	300	360
c	210	15	400	460
d	4	10	600	660

فإذا اعتبرنا سنة 2002 هي سنة الأساس، فأوجد ما يلي:

a) رقم لاسبير للأسعار.

b) رقم باش للأسعار.

c) رقم فشر للأسعار.

الحل:

a) $$laspeyre(P) = \frac{\sum P_1Q_0}{\sum P_0Q_0} \times 100\%$$

$$= \frac{40\times 200+20\times 300+15\times 400+10\times 600}{28\times 200+16\times 300+210\times 400+4\times 600}$$

$$= \frac{26000}{96800}\times 100\%$$

$$= 26.86\%$$

b) $$\boldsymbol{paasche(P)} = \frac{\sum P_1 Q_1}{\sum P_0 Q_1}\times 100\%$$

$$= \frac{40\times 250+20\times 360+15\times 460+10\times 660}{28\times 250+16\times 360+210\times 460+4\times 660}\times 100\%$$

$$= \frac{30700}{112000}\times 100\%$$

$$= 27.41\%$$

$$Fisher(p) = \sqrt{Laspeyre(P)\times Paasche(P)}\ \%$$

$$= \sqrt{26.86\times 27.41}\ \%$$

$$= 27.13\%$$

ملاحظة: بعض الباحثين الإحصائيين يستخدمون أوزان تعطى للأسعار بدل الكميات، ولكننا في هذا الكتاب سنستخدم الكميات مباشرة وليس أوزان الأسعار.

2- الرقم القياسي المرجح للكميات: وهو ثلاثة أرقام.

a- الرقم القياسي للكميات المرجح بأسعار سنة الأساس (رقم لاسبير للكميات)

رقم لاسبير للكميات= $\frac{\text{مجموع (كميات سنة المقارنة× أسعار سنة الأساس)}}{\text{مجموع (كميات سنة الأساس× أسعار سنة الأساس)}}$ × 100٪

$$\boldsymbol{laspeyre(Q)} = \frac{\sum Q_1 P_0}{\sum Q_0 P_0}\times 100\%$$

b- الرقم القياسي للكميات والمرجح بأسعار سنة المقارنة (رقم باش للكميات).

رقم باش للأسعار= $\frac{\text{مجموع (كميات سنة المقارنة× أسعار سنة المقارنة)}}{\text{مجموع (كميات سنة الأساس× أسعار سنة المقارنة)}}$ × 100٪

$$paasche(Q) = \frac{\sum Q_1 P_1}{\sum Q_0 P_1} \times 100\%$$

c- الرقم القياسي الأمثل للكميات: (رقم فشر للكميات).

$$Fisher(Q) = \sqrt{Laspeyre \times Paasche}\ \%$$

مثال:

الجدول التالي يبين أسعار وكميات أربع سلع في عامي 2001,2002.

السلعة	سعر الوحدة		كمية المبيعات	
	عام 2001	عام 2002	عام 2001	عام 2002
a	20	25	200	240
b	22	19	180	210
c	20	20	300	280
d	9	11	110	100

على اعتبار أن سنة 2001 هي سنة الأساس أوجد ما يلي:

a) رقم لاسبير للكميات.

b) رقم باش للكميات.

c) رقم فشر للكميات.

الحل

a) $$laspeyre(Q) = \frac{\sum Q_1 P_0}{\sum Q_0 P_0} \times 100\%$$

$$= \frac{240 \times 20 + 210 \times 22 + 280 \times 20 + 100 \times 9}{200 \times 20 + 180 \times 22 + 300 \times 20 + 110 \times 9} 100\%$$

$$= \frac{15920}{14950} \times 100\%$$

$$= 106.49\%$$

b) $$Paasche(Q) = \frac{\sum Q_1 P_1}{\sum Q_0 P_1} \times 100\%$$

$$= \frac{240 \times 25 + 210 \times 19 + 280 \times 20 + 100 \times 11}{200 \times 25 + 180 \times 19 + 300 \times 20 + 110 \times 11} \times 100\%$$

$$= \frac{16690}{15630} \times 100\%$$

$$= 106.78\%$$

c) $$Fisher(Q) = \sqrt{Laspeyre \times Paasche}\ \%$$

$$= \sqrt{106.49 \times 106.78}\ \%$$

$$= 106.63\%$$

ملاحظة: إذا لم تذكر سنة الأساس لظاهر ما. فتكون السنة الأقدم هي سنة الأساس.

رقم مارشال:

هنالك عالم آخر اسمه مارشال يرجح الأرقام القياسية للأسعار (للكميات) بمعدل كميات (أسعار) سنة المقارنة وسنة الأساس فيكون:

a) $$\boldsymbol{Marshall(P) = \frac{\sum P_1(Q_0 + Q_1)}{\sum P_0(Q_0 + Q_1)} \times 100\%}$$

b) $$\boldsymbol{Marshall(Q) = \frac{\sum Q_1(P_1 + P_0)}{\sum Q_0(P_1 + P_0)} \times 100\%}$$

ففي المثال السابق يكون:

a) $$\text{Marshall(p)} = \frac{25(200+240)+19(180+210)+20(300+280)+11(110+100)}{20(200+240)+22(180+210)+20(300+280)+9(110+100)} \times 100\%$$

$$= 104.7\%$$

b) $$\text{Marshall(Q)} = \frac{240(20+25)+210(22+19)+280(20+20)+100(9+11)}{200(20+25)+180(22+19)+300(20+20)+110(9+11)} \times 100\%$$

$$= 106.64\%$$

تمارين

1- في مصنع للمعلبات ينتج خمسة أنواع من المعلبات كانت أسعار وكميات عامي 1999,2000كالآتي:

السلعة	سعر الوحدة بالدينار		الكمية بالطن	
	عام1999	عام 2000	عام 1999	عام 2000
فول	0.20	0.23	150	200
حمص	0.25	0.27	100	250
بندورة	0.12	0.15	110	300
فاصولياء	0.30	0.35	120	350
بازيلاء	0.25	0.30	130	220

أوجد ما يلي باعتبار سنة 1999 هي سنة الأساس:

a) الرقم القياسي التجميعي البسيط للأسعار.

b) الرقم القياسي النسبي البسيط للأسعار.

c) الرقم القياسي التجميعي البسيط للكميات.

d) الرقم القياسي النسبي البسيط للكميات.

2- فيما يلي جدول يبين أسعار وكميات مبيعات مجموعة من السلع في عامي 2003 ,2000.

السلعة	سعر الوحدة بالدينار		كمية المبيعات	
	عام 2000	عام 2003	عام 2000	عام 2003
a	40	52	300	250
b	28	32	400	480
c	21	27	500	560
d	16	18	700	770
e	36	232	800	1000

احسب ما يلي باعتبار عام 2000 هو عام الأساس:

a) رقم لاسبير للأسعار.

b) رقم باش للأسعار.

c) رقم فشر للأسعار.

d) رقم مارشال للأسعار.

3- الجدول التالي يمثل أسعار وكميات خمس سلع في عامي1999,2002.

السلعة	سعر الوحدة بالدينار		الكمية	
	عام 1999	عام 2002	عام 1999	عام 2002
a	10	20	30	40
b	18	20	22	24
c	40	45	50	55
d	10	11	12	13
e	9	13	17	19

احسب ما يلي باعتبار عام 1999 هي سنة الأساس:

a) رقم لاسبير للكميات.

b) رقم باش للكميات.

c) رقم فشر للكميات.

d) رقم مارشال للكميات.

4- إذا كان لدينا سلعتين كالعدس والقمح. وكان سعر العدس في سنة المقارنة ضعف ما كان عليه في سنة الأساس والكمية المستهلكة من العدس نصف ما كانت عليه في سنة الأساس، ولم يطرأ تغير على السعر والكمية ما بين سنة المقارنة والأساس للقمح. فما هو الرقم القياسي للأسعار والمرجح بالكميات؟

5- ما هو الرقم القياسي لسلعة ما في سنة 2005 مقارنة بنفس السنة؟

6- الجدول التالي يبين أسعار وكميات السيارات المباعة لدى تاجر سيارات في الفترة 1995-1997:

السلعة	سعر السيارة بالدينار			عدد السيارات المباعة		
	1995	1996	1997	1995	1996	1997
a	7000	8000	8200	20	30	45
b	10000	14000	14000	40	35	50
c	25000	32000	34000	15	20	30

احسب باعتبار عامي 1995,1996 هما الأساس:

a) الرقم القياسي التجميعي البسيط للأسعار والكميات.

b) الرقم القياسي النسبي البسيط للأسعار والكميات.

c) رقم لاسبير للأسعار.

d) رقم باش للأسعار.

e) رقم فشر للأسعار.

تحسب قيمة سنة الأساس هنا على أنها الوسط الحسابي لسنتي 1995,1996.

7- إذا كان رقم لاسبير للأسعار يساوي 110%. ورقم فشر للأسعار يساوي 115%. أوجد رقم باش للأسعار؟

8- إذا كانت الأجور اليومية لثلاثة عمال سنة 2004 هي بالدينار 6,5,7وكانت الأرقام القياسية المقابلة للأجور اليومية لهؤلاء العمال 120,100,87.5 على الترتيب باعتبار سنة 2003 هي الأساس. جد معدل الأجور اليومية لهؤلاء العمال سنة 2003؟

9- إذا كان معدل تكاليف المعيشة ومعدل دخل الفرد في عام 1995 هو (100) ديناراً، (180) ديناراً شهرياً على الترتيب وأصبح في عام 2000(350) ديناراً، (250) ديناراً. فما هي القوى الشرائية لدخل الفرد في عام 2000 باعتبار عام 1995 هو الأساس. (فسر إجابتك).

10- الجدول التالي يمثل معدل أسعار وعدد الأسهم المباعة لأربع شركات بين شهري آب وتموز من عام 1997.

الشركة	سعر السهم بالدينار		عدد الأسهم المباعة	
	آب	تموز	آب	تموز
الأردنية	1	1.25	1000	500
الوطنية	x	0.8	2000	1500
العربية	2	2.5	200	300
الدولية	1.5	y	1300	1300

إذا كـان رقـم لاسـبير للأسـعار = 100.95٪، ورقـم بـاش للأسـعار = 101.12٪ أوجد سعر سهم الـشركة الوطنيـة في شـهر آب مـن عـام 1997، وسعر سهم الشركة الدولية في شهر تموز من عام 1997.



الوحدة التاسعة

السلاسل الزمنية

Time Series

السلاسل الزمنية
Time Series

مقدِّمة

إذا أخذنا كميات المطر في أحد أشهر الشتاء لعدة سنوات متتالية فإن هذه الكميات تشكل سلسلة زمنية، ومن هذه السلسلة الزمنية يمكن التنبؤ بكمية المطر في ذلك الشهر لسنوات لاحقة بناء على بيانات السنوات السابقة.

تعريف:

السلسلة الزمنية هي مجموعة مشاهدات أخذت على فترات زمنية متلاحقة ويفضل تساوي الفترات الزمنية التي تأخذ فيها المشاهدات.

من الأمثلة على السلاسل الزمنية: أخذ كمية الفوسفات التي يصدرها الأردن سنويا في سنوات متتالية، مبيعات أحد المتاجر لمدة عشرة أعوام متتالية.

من استعمالات السلاسل الزمنية:

a) التنبؤ بالمستقبل باستعمال البيانات الإحصائية التي أخذت في الماضي.

b) اكتشافات الدورات التي تتكرر فيها البيانات.

c) اكتشاف الطفرات الاقتصادية التي تحصل في زمن ما.

تمثيل السلسلة الزمنية بيانياً Graphs of time series

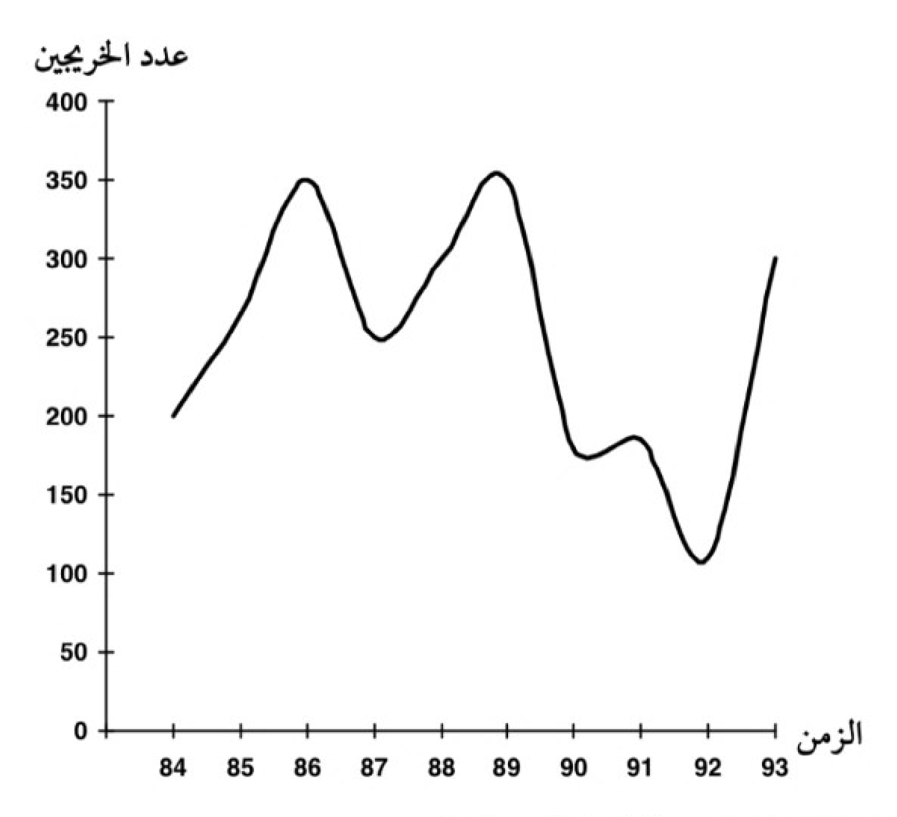
يمكن تمثيل السلسلة الزمنية بيانياً بتعيين أزواج مرتبة (الزمن، قيمة الظاهرة) في المستوى البياني ثم توصيل تلك النقاط. ويسمى المنحنى الناتج المنحنى التاريخي (Historical curve) للسلسلة الزمنية.

مثال:

ارسم المنحنى التاريخي الذي يمثل السلسلة الزمنية لعـدد خريجـي إحـدى كليـات المجتمع خلال السنوات 1984-1993.

السنة	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
عدد الخريجين	200	265	350	250	300	350	180	185	110	300

الحل:



معامل الخشونة والمتوسطات المتحركة

Roughness Coefficient and moving average

إذا نظرنا إلى المنحنى التاريخي للسلسلة الزمنية في المثال السابق لأعـداد الخـريجين من إحدى كليات المجتمع، نرى أنها ترتفـع في بعـض الـسنوات وتـنخفض في سـنوات أخرى وهذا التذبذب يسمى خشونة السلسلة الزمنية ولحساب خـشونة سلـسلة زمنيـة ما نستخدم مقياس يسمى معامل الخشونة (R.C):

$$R.C = \frac{\sum_{i=2}^{n}(x_i - x_{i-1})^2}{\sum_{i=2}^{n}(x_i - \bar{x})^2}$$

حيث x_i: المشاهدة رقم(i) في السلسلة الزمنية.

وكلما قل معامل الخشونة كانت السلسلة ملساء أكثر.

مثال:

احسب معامل الخشونة للسلسلة.

9, 0, 3, 9, 6, 3, 9, 0, 6

الحل:

في البداية نجد الوسط الحسابي للسلسلة

$$\bar{x} = \frac{6+0+9+3+6+9+3+0+9}{9} = 5$$

$$\sum_{i=2}^{9}(x_i - x_{i-1})^2 = (0-6)^2 + (9-0)^2 + (3-9)^2 + (6-3)^2 + (9-6)^2$$
$$+ (3-9)^2 + (0-3)^2 + +(9-0)^2$$
$$= 36+81+36+9+9+36+9+81$$
$$= 297$$

$$\sum_{i=2}^{9}(x_i - \bar{x})^2 = (0-5)^2 + (9-5)^2 + (3-5)^2 + (6-5)^2$$
$$(9-5)^2 + (3-5)^2 + (0-5)^2 + (9-5)^2$$
$$= 25 + 16 + 4 + 1 + 16 + 4 + 25 + 16$$
$$= 107$$

$$R.C = \frac{297}{107} = 2.78$$

نرى هنا أن معامل الخشونة كبير نسبياً ولـذلك تكـون الدراسـة الإحـصائية الـتي يمكن أن تجرى على هذه السلسلة غير دقيقة النتائج وسيكون تحليلها صعب نوعا ما.

ولذلك لا بد من تقليل معامل الخشونة وذلك عن طريق إيجاد معدلات متحركـة بطول محدد لتكون سلسلة أخرى أقل تذبذبا.

فإذا أردنا إيجاد معدلات متحركة بطول (m) سنرمز له بالرمز $\overline{M}av(m)$ فإن هـذا المعدل يكون

$$\overline{M}av(m) = \frac{Xr + X_{r+1} + \ldots + X_{r+(m-1)}}{m} \qquad r = 1,2,3,\ldots$$

فمثلاً إذا أردنا إيجاد معدل متحرك للسلسلة السابقة بطول (3) فإن هذا المعدل يكون

$$\overline{M}av(3) = \frac{X_r + X_{r+1} + X_{r+2}}{3}$$

وبالتالي نجد عناصر السلسلة الجديدة وهي

$$\frac{X_1 + X_2 + X_3}{3} = \frac{6+0+9}{3} = \frac{15}{3} = 5$$

$$\frac{X_2 + X_3 + X_4}{3} = \frac{0+9+3}{3} = \frac{12}{3} = 4$$

$$\frac{X_3 + X_4 + X_5}{3} = \frac{9+3+6}{3} = \frac{18}{3} = 6$$

$$\frac{X_4 + X_5 + X_6}{3} = \frac{3+6+9}{3} = \frac{18}{3} = 6$$

$$\frac{X_5 + X_6 + X_7}{3} = \frac{6+9+3}{3} = \frac{18}{3} = 6$$

$$\frac{X_6 + X_7 + X_8}{3} = \frac{9+3+0}{3} = \frac{12}{3} = 4$$

$$\frac{X_7 + X_8 + X_9}{3} = \frac{3+0+9}{3} = \frac{12}{3} = 4$$

وبالتالي تصبح السلسلة الجديدة هي:

5, 4, 6, 6, 6, 4, 4

ولحساب معامل الخشونة لهذه السلسلة سيكون وسطها الحسابي هو:

$$\overline{X} = \frac{5+4+6+6+6+4+4}{7} = \frac{35}{7} = 5$$

ويكون معامل الخشونة R.C هو:

$$R.C = \frac{\sum_{i=2}^{7}\left(x_i - x_{i-1}\right)^2}{\sum_{i=2}^{7}\left(x_i - \overline{x}\right)^2}$$

$$\begin{aligned}\sum_{i=2}^{7}\left(x_i - x_{i-1}\right)^2 &= (4-5)^2 + (6-4)^2 + (6-6)^2 + (6-6)^2 \\ &\quad + (4-6)^2 + (4-4)^2 \\ &= 1+4+0+0+4+0 \\ &= 9\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{7}\left(x_i - \overline{x}\right)^2 &= (5-5)^2 + (4-5)^2 + (6-5)^2 + (6-5)^2 + (6-5)^2 \\ &\quad + (4-5)^2 + (4-5)^2 \\ &= 0+1+1+1+1+1+1 \\ &= 6\end{aligned}$$

$$\therefore R.C = \frac{9}{6} = 1.5$$

ملاحظة: عدد الأوساط المتحركة بطول (m) هو k=n-(m-1)

حيث n عدد مفردات السلسلة الأصلية

K عدد الأوساط المتحركة

مركبات السلاسل الزمنية

السلسلة الزمنية تتكون من أربع مركبات هي:

1- مركبة الاتجاه (Secular Trend): وتمثل الاتجاه للسلسلة ويكون التقدير الأفضل لها عن طريق معادلة خط انحدار قيمة الظاهرة x على الزمن t.

2- مركبة الدورة (Cyclical movements): تمثل فترة تغير البيانات لمدة طويلة قد تزيد عن السنة.

3- المركبة الفصلية (Seasonal movements): وهي التغيرات المنتظمة التي تظهر في الفصول والفصول قد تكون ربع سنوية أو شهرية أو أسبوعية.

4- مركبة الخطأ (أو المركبة العشوائية)(Irregular or random movements) : تصف ما تبقى من العوامل التي لم تدخل في المركبات السابقة كالرواج الاقتصادي غير المتوقع في إحدى سنوات السلسلة الزمنية أو الركود نتيجة لكوارث.

ويقصد بتحليل السلسلة الزمنية هو إظهار تأثير إحدى المركبات السابقة بعد إلغاء تأثير المركبات الأخرى.

تقدير مركبة الاتجاه (Estimation of secular trend)

تقدر مركبة الاتجاه بعدة طرق منها:

a- طريقة المربعات الصغرى (Least squares method)

وهي إيجاد معادلة خط انحدار قيمة الظاهرة (x) على الزمن (t) وتسمى معادلة الاتجاه العام (Equation of trend).

وهي

$$x = a\,t + b$$

حيث

$$a = \frac{\sum_{i=1}^{k} x_i t_i - k\bar{x}\bar{t}}{\sum_{i=1}^{k} t_i^2 - k(\bar{t})^2}$$

$$b = \bar{x} - a\,\bar{t}$$

قيمة الظاهرة :x

الزمن : t

عدد قيم الظاهرة :k

وبعد إيجاد معادلة الاتجاه العام يمكن تقدير قيم الظاهرة عن طريقها.

مثال:

الجدول التالي يمثل إنتاج أحد المصانع بآلاف الدنانير في الفترة (1980-1989):

السنة	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
الإنتاج	200	235	195	210	245	200	220	260	230	245

أوجد معادلة الاتجاه العام ثم اكتب الإنتاج التقديري للمصنع في جميع السنوات؟

الحل:

لسهولة التعامل نعطي السنوات ترقيم من 1 إلى 10 (أي نطرح 1979 من كل سنة) ونجد معادلة الاتجاه العام

$$x = a\,t + b$$

الزمن t	الإنتاج x	xt	t^2
1	200	200	1
2	235	470	4
3	195	585	9
4	210	840	16
5	245	1225	25
6	200	1200	36
7	220	1540	49
8	260	2080	64
9	230	2070	81
10	245	2450	100
55	2240	12660	385

$\bar{x} = 224$, $\bar{t} = 5.5$

$$a = \frac{\sum_{i=1}^{k} x_i t_i - k\bar{x}\bar{t}}{\sum_{i=1}^{k} t_i^2 - k(\bar{t})^2}$$

$$= \frac{12660 - 10(224)(5.5)}{385 - 10(5.5)^2}$$

$$= 4.12$$

$$b = \bar{x} - a\bar{t}$$

$$= 224-(4.12)(5.5)$$

$$= 201.34$$

وتكون معادلة الاتجاه العام هي:

$$x = 4.12(t) + 201.34$$

ولتقدير إنتاج المصنع نطبق في معادلة الاتجاه العام.

فمثلاً عندما $t = 1$ فإن $x = 4.12(1) + 201.34$

$= 205.46$

عندما $t=2$ فإن $x = 4.12(2) + 201.34$

$= 209.58$

ونستمر هكذا حتى نحصل على كل القيم المقرة لـ x فنحصل على العمود التالي:

قيمة x المقدرة	الزمن t
205.46	1
209.58	2
213.7	3
217.82	4
221.94	5
226.06	6
230.18	7
234.3	8
238.42	9
242.54	10

مثال:

الجدول التالي يمثل الإنتاج الصناعي لإحدى الدول بملايين الدولارات في الفترة (1974-1980)

السنة	1974	1975	1976	1977	1978	1979	1980
الإنتاج الصناعي	101	95	104	108	106	111	117

1- أوجد معادلة الاتجاه العام.

2- قدر الإنتاج الصناعي لكل سنة.

3- ما هو الإنتاج المتوقع للدولة سنة 1985.

الحل:

الزمن t	الإنتاج x	tx	t^2
1	101	101	1
2	95	190	4
3	104	312	9
4	108	432	16
5	106	530	25
6	111	666	36
7	117	819	49
28	742	3050	140

1) $\bar{x} = \frac{742}{7} = 106$

$$\bar{t} = \frac{28}{7} = 4$$

$$a = \frac{\sum xt - k\bar{x}\bar{t}}{\sum t^2 - k(\bar{t})^2}$$

$$= \frac{3050 - (7)(106)(4)}{140 - 7(4)^2} = 2.93$$

$$b = \bar{x} - a\bar{t}$$

$$=106 - (2.93)\ (4)$$

$$= 94.3$$

∴ معادلة خط الانحدار (معادلة الاتجاه العام) هي

$$x = 2.93t+94.3$$

2- لإيجاد الإنتاج الصناعي المقدر نعوض في المعادلة السابقة للحصول على الجدول:

السنة	الزمن t	الإنتاج الصناعي المقدر
1974	1	97.23
1975	2	100.16
1976	3	103.09
1977	4	106.02
1978	5	108.95
1979	6	111.88
1980	7	114.81

3- سنة 1985 تقابل الترتيب t =12

فيقدر الإنتاج

$$x =(2.93)\ (12)+ 94.3$$
$$= 129.43 \text{ million } \$$$

b- طريقة التمهيد باليد (Free hand method)

وتتم هذه الطريقة برسم خط مستقيم متوافق مع نقاط المنحنى التاريخي للسلسلة الزمنية، وهي طريقة تعتمد على مهارة الذي يرسم ذلك المستقيم، ولذلك تعتبر طريقة غير دقيقة.

وبعد رسم المستقيم نجد معادلته عن طريق نقطتين عليه.

وتكون معادلة المستقيم هي معادلة الاتجاه العام.

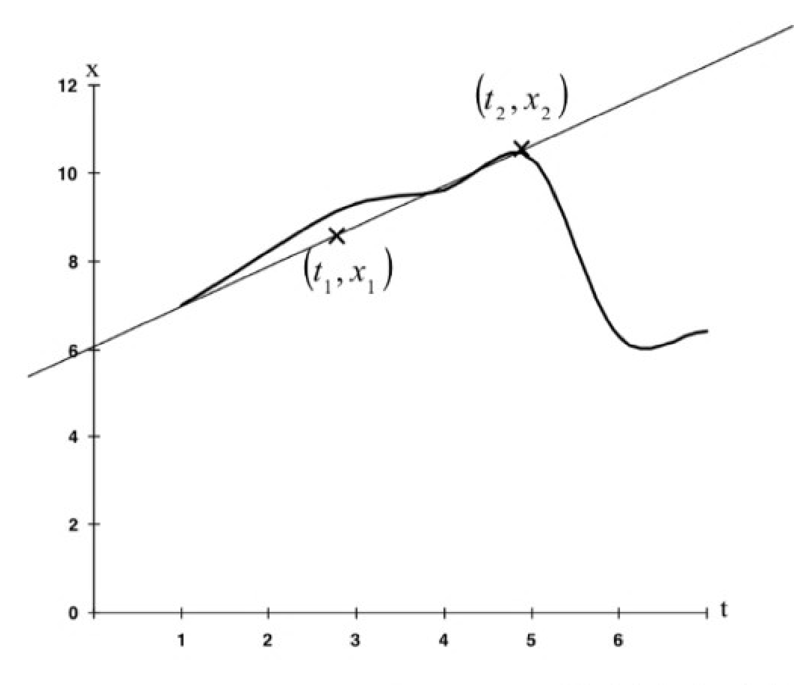
مثال:

الجدول التالي يمثل ميزانية التعليم العالي بملايين الدنانير للفترة (1992-1986)

الزمن t	موازنة الوزارة	السنة
1	7	1986
2	8.2	1987
3	9.3	1988
4	9.6	1989
5	10.3	1990
6	6.3	1991
7	6.4	1992

a) جد معادلة الاتجاه؟

b) قدر ميزانية الوزارة في كل سنة من سنوات الجدول؟



نجد في البداية ميل الخط المستقيم وهو:

$$m = \frac{x_2 - x_1}{t_2 - t_1} = \frac{10-8}{4.5-2.3} = \frac{2}{2.2} = 0.91$$

وتكون معادلة خط المستقيم هي:

$$(x-x_1) = m\,(t-t_1)$$
$$(x-8) = 0.91(t-2.3)$$
$$x-8 = 0.91t-2.09$$

$\therefore x=0.91t + 5.91$ (معادلة الاتجاه العام)

السنة	1986	1987	1988	1989	1990	1991	1992
الميزانية المقدرة	6.82	7.73	8.64	9.55	10.46	11.37	12.28

C– طريقة نصف السلسلة(Semi-averages method)

وتتم هذه الطريقة بقسم السلسلة إلى نصفين متساويين نجد الوسط الحسابي للنصف الأول بحيث تشكل نقطة في المستوى البياني، الإحداثي الأفقي لها هو الوسط

الحسابي لقيم الزمن $(\bar{t}_1)$ والإحداثي العمـودي هـو الوسـط الحـسابي لقـيم الظـاهرة $(\bar{x}_1)$ في ذلك النصف، ونقوم بنفس العملية للنصف الآخر فنحصل على نقطـتين همـا $(\bar{t}_1, \bar{x}_1), (\bar{t}_2, \bar{x}_2)$ ثم نجد معادلة الخط المستقيم المـار بـالنقطتين الـسابقتين وتكـون هـي معادلة الاتجاه العام.

مثال:

الجدول التالي يمثل أعداد الطلبة بآلاف والـذين يدرسـون خـارج الأردن خـلال الفترة (1988-1993).

السنة	1988	1989	1990	1991	1992	1993
عدد الطلبة بالآلاف	36	35.6	32.8	35.8	35	31.9

الحل:

السنة	أعداد الطلبة	الزمن t	متوسط الزمن t	متوسط أعداد الطلبة x
1988	36	1	$\bar{t}_1 = 2$	$\bar{x}_1 = 34.8$
1989	35.6	2		
1990	32.8	3		
1991	35.8	4	$\bar{t}_2 = 5$	$\bar{x}_2 = 34.23$
1992	35	5		
1993	31.9	6		

نجد ميل الخط المستقيم

$$m = \frac{\bar{x}_2 - \bar{x}_1}{\bar{t}_2 - \bar{t}_1} = \frac{34.23 - 34.8}{5 - 2}$$

$$= -0.19$$

معادلة الخط المستقيم

$$x - \bar{x}_1 = m(t - \bar{t}_1)$$

$$x - 34.8 = -0.19\ (t-2)$$

$$x - 34.8 = -0.19t + 0.38$$

$$x = -0.19t + 35.18$$

وتكون معادلة الاتجاه العام هي: $x= 35.18 - 0.19t$

سؤال: في المثال السابق اكتب أعداد الطلبة المقدرة لكل سنة من السنوات 1988-1993.

ملاحظة:

إذا كان عدد عناصر السلسلة فردي فإننا نحذف القيمة الموجودة في منتصف السلسلة ونأخذ الأوساط للقيم المتبقية، فمثلاً إذا كان عدد عناصر السلسلة (9) فإننا نأخذ المتوسط لأول أربع قيم والمتوسط لآخر أربع قيم ونحذف القيمة الخامسة.

d- طريقة المتوسطات المتحركة (Moving averages method)

تتلخص هذه الطريقة بإيجاد الأوساط (المتوسطات) المتحركة بطول مناسب للسلسلة الزمنية فينتج لدينا سلسلة زمنية أخرى من المتوسطات المتحركة.

ويكون أثر الاتجاه العام فيها ظاهراً بشكل أفضل من السلسلة الزمنية الأصلية. ثم نقدر الاتجاه العام بإحدى الطرق آنفة الذكر.

مثال:

الجدول التالي يمثل أعداد الطلبة بآلاف والمسجلين في إحدى كليات المجتمع خلال الفترة 1985-1994.

السنة	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
عدد الطلاب بآلاف	2	3	1	4	6	3	2	3	1	2

1- ارسم المنحنى التاريخي للسلسلة الزمنية.

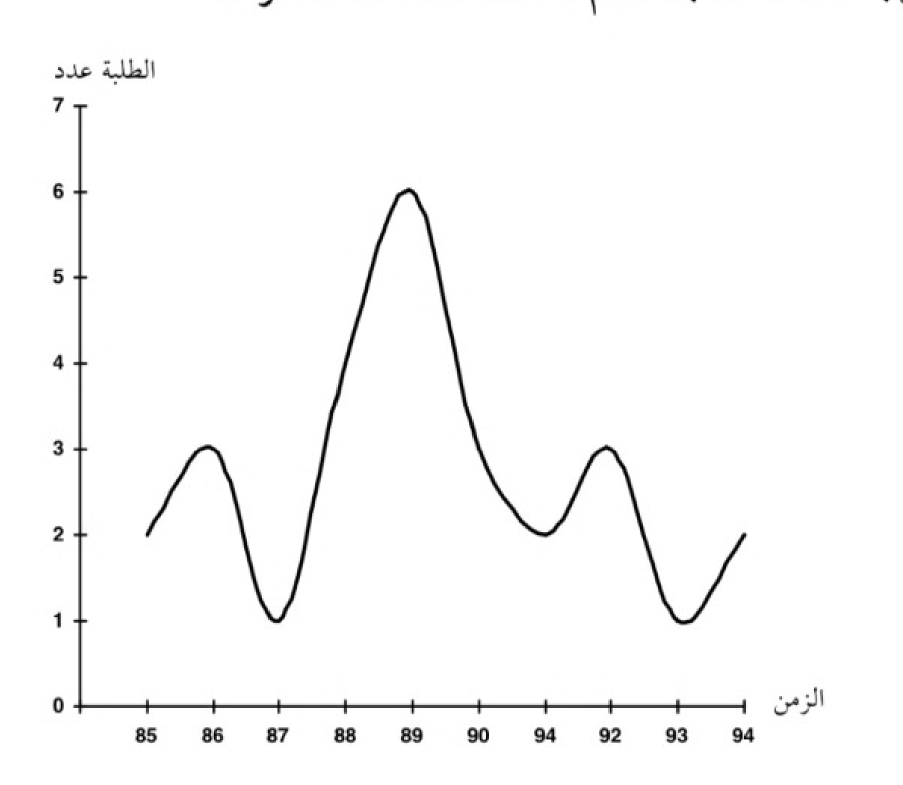
2- أوجد معادلة الاتجاه العام للسلسلة الزمنية.

3- أوجد سلسلة المعدلات المتحركة بطول(3). وأرسم المنحنى التاريخي لها.

4- أوجد معادلة الاتجاه العام لسلسلة المعدلات المتحركة.

الحل:

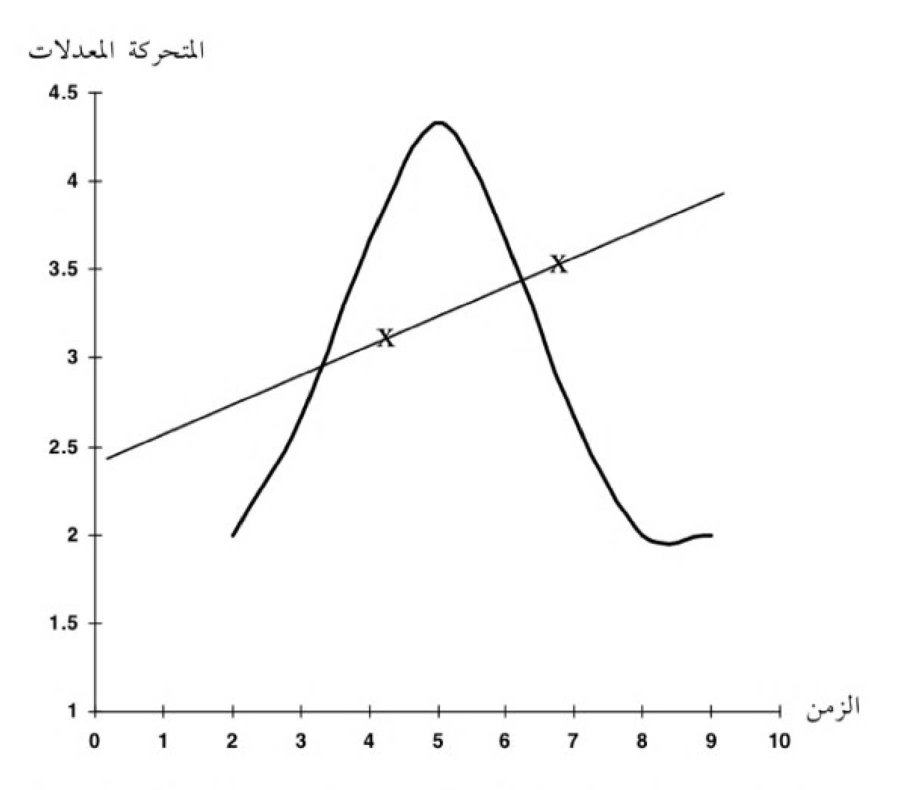
-1



2- تمرين.

3- نجد المعادلات المتحركة بطول (3) للسلسلة الزمنية.

المعدلات المتحركة	عدد الطلبة بالآلاف	السنة	الزمن t
-	2	1985	1
2	3	1986	2
2.67	1	1987	3
3.67	4	1988	4
4.33	6	1989	5
3.67	3	1990	6
2.67	2	1991	7
2	3	1992	8
2	1	1993	9
-	2	1994	10



4- نجد معادلة الاتجاه العام بطريقة التمهيد باليد ويكون أفضل خط مستقيم "يمر بالنقطتين (3.5,7) , (4,3) " لذلك تكون معادلة الاتجاه العام هي:

$$\frac{x-3}{t-4}=\frac{3.5-3}{7-4}$$

$$\frac{x-3}{t-4}=\frac{0.5}{3}$$

$$\frac{x-3}{t-4}=\frac{1}{6}$$

بالضرب التبادلي ينتج أن

$$6x-18 \quad = t-4$$

$$6x \ = t+14$$

$$\therefore \qquad x=\frac{1}{6}(t+14)$$

سؤال: جد معادلة الاتجاه لسلسلة المتوسطات المتحركة بطريقة أخرى؟

ملاحظة: عند إيجاد سلسلة المتوسطات المتحركة بطول m يكون أول وسط متحرك مقابل وسيط أول m من الأزمنة.

ففي مثالنا السابق لو كان الطـول المتحـرك 4 فيكـون الوسـط الحـسابي المتحـرك الأول مقابل وسيط السنوات 1985، 1986، 1987، 1988 أي منتصف عام 86.

تقدير المركبة الفصلية: Estimating Seceonal movement

هنالك أربع طرق لتقدير مركبة الفصل وهي:

أ- طريقة النسبة المئوية للمعدل (Average percentage method)

ب- طريقة النسبة إلى الاتجاه العام (Ratio to trend method)

ج- طريقة النسبة إلى المتوسطات المتحركة (Ration to moving averages method)

د- طريقة الوصل النسبية (the link relative method)

وسنكتفي في هذا الكتاب بعرض الطريقة الأولى والتي سنوضحها في المثال التالي:

مثال:

الجدول التالي يبين المبيعـات الـشهرية لأحـد المتـاجر بمئـات الـدنانير للـسنوات (1993-1997):

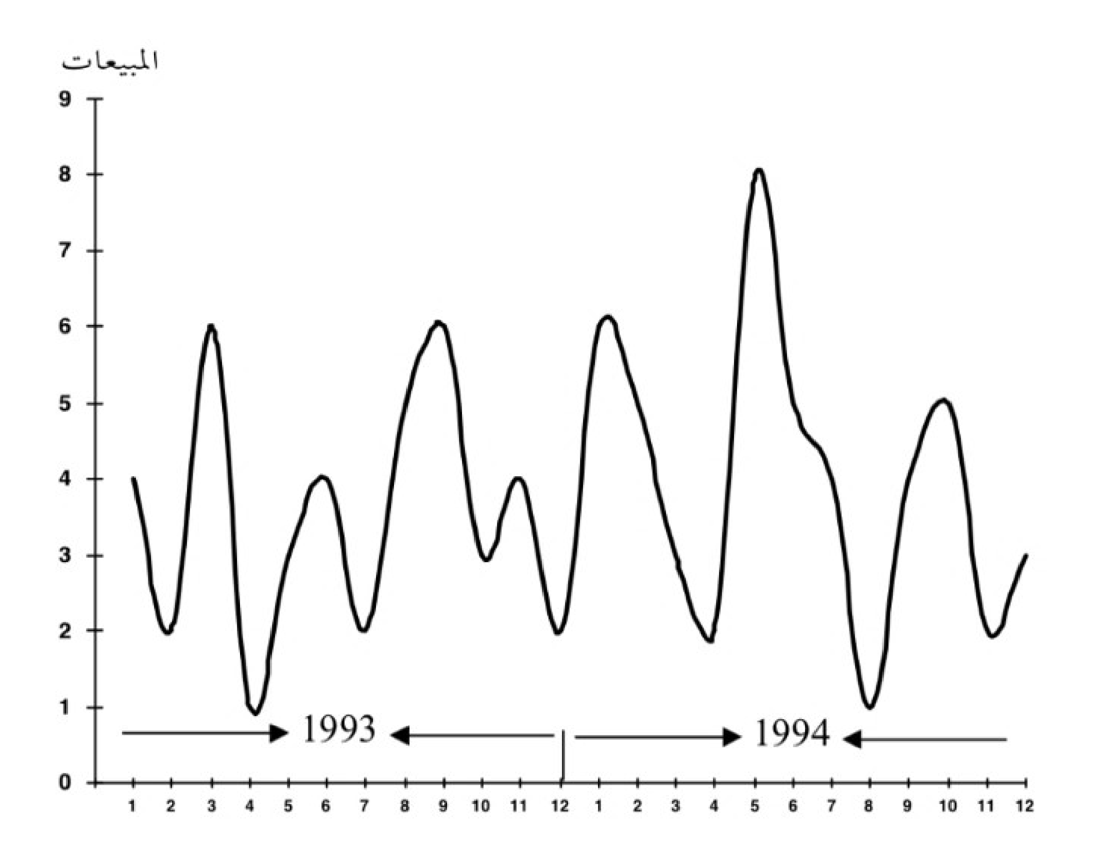
الشهر / السنة	1	2	3	4	5	6	7	8	9	10	11	12
1993	4	2	6	1	3	4	2	5	6	3	4	2
1994	6	5	3	2	8	5	4	1	4	5	2	3
1995	5	3	2	4	5	4	3	1	2	4	1	2
1996	7	6	4	3	7	6	5	4	5	6	3	4
1997	3	2	5	4	1	4	5	8	2	3	5	6

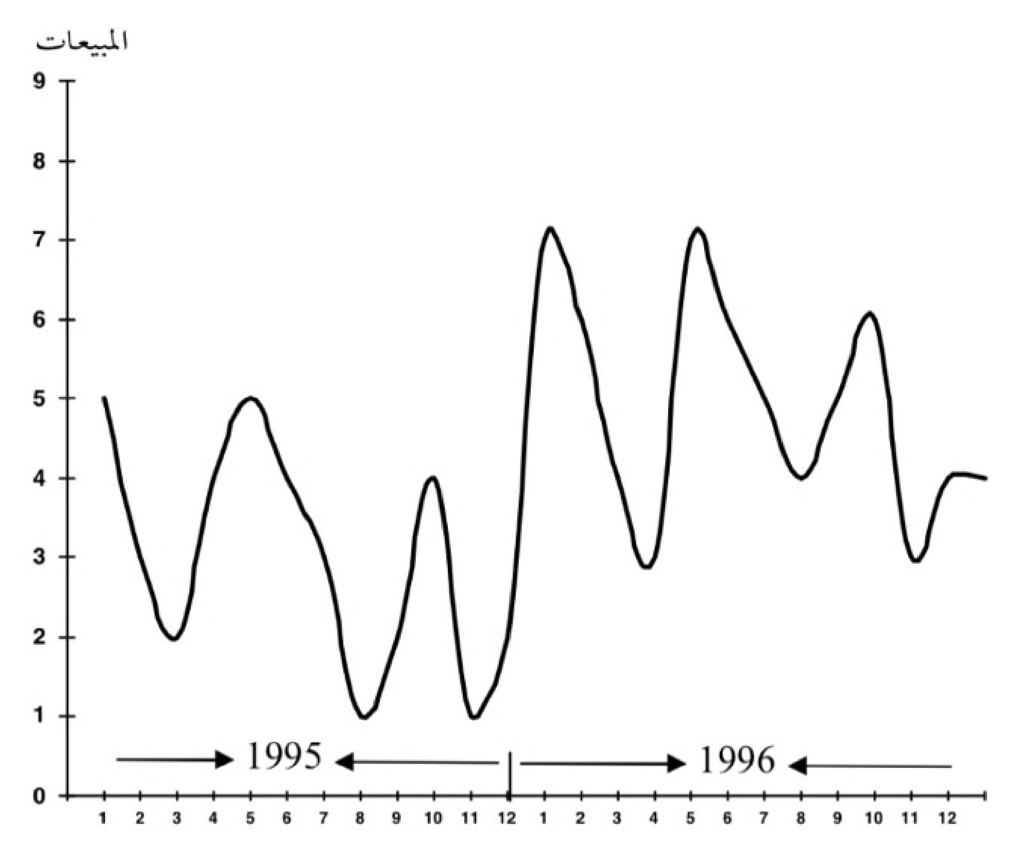
a) أرسم المنحنى التاريخي للسلسلة.

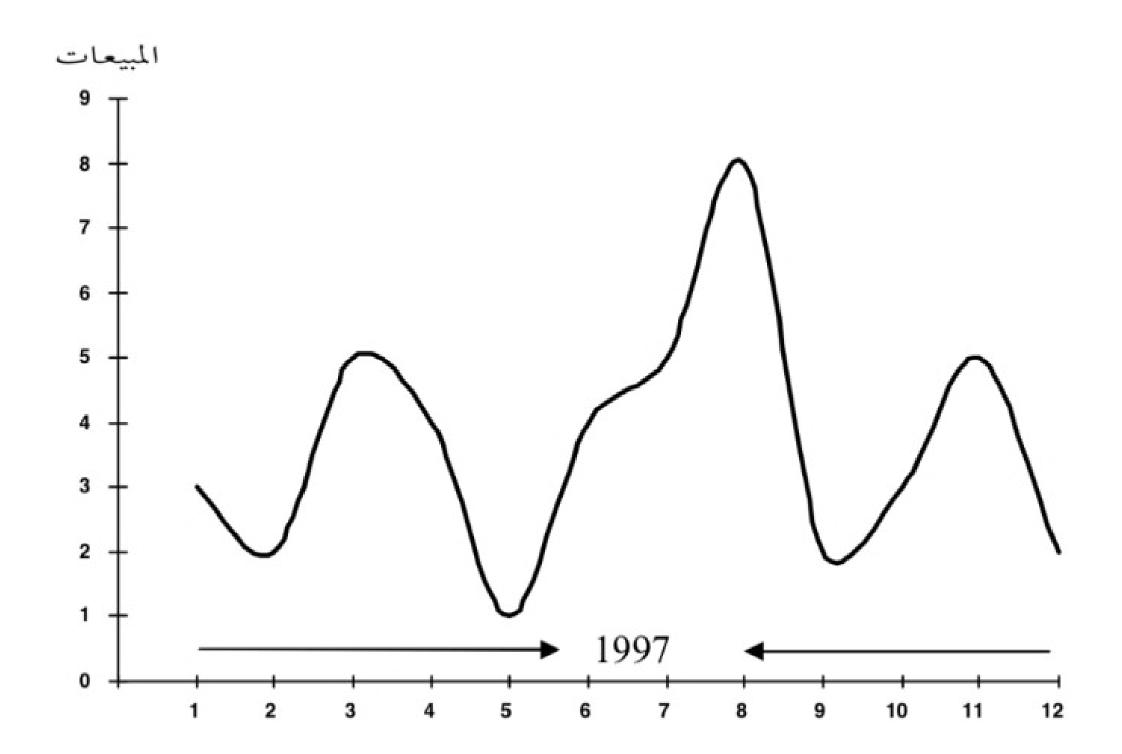
b) حلل مركبة الفصل بطريقة النسبة المئوية للمعدل.

الحل:

a- (لاحظ أن الفصل هنا عبارة عن شهر)


المبيعات
9
8
7
6
5
4
3
2
1
0
1993
1994
1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12


المبيعات
9
8
7
6
5
4
3
2
1
0
1995
1996
1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12



b- **أولاً: نجد المعدل الشهري (الفصلي) للمبيعات** لكل سنة على حدة والجدول التالي يبين مجموع المبيعات السنوي وكذلك معدل المبيعات الشهري لكل سنة:

المعدلات الشهرية	المجموع	السنة
3.5	42	1993
4	48	1994
3	36	1995
5	60	1996
4	48	1997

ثانياً: نجد الرقم القياسي الفصلي ويعطى بالعلاقة

$$\text{الرقم القياسي الفصلي} = \frac{\text{قيمة الظاهرة في سنة ما}}{\text{المعدل الفصلي لتلك السنة}} \times 100\%$$

وفي مثالنا هذا يكون:

$$\text{الرقم القياسي الشهري} = \frac{\text{قيمة المبيعات الشهرية في سنة ما}}{\text{المعدل الشهري لتلك السنة}} \times 100\%$$

وبتطبيق العلاقة على شهر (1) من سنة 1993 يكون:

الرقم القياسي الشهري = $\frac{4}{3.5} \times 100\% = 114.3\%$

والجدول التالي يعطي الرقم القياسي للأشهر المختلفة لكل سنة:

الشهر	السنة				
	1993	1994	1995	1996	1997
1	114.3	150	166.7	140	75
2	57.1	125	100	120	50
3	171.4	75	66.7	80	125
4	28.6	50	133.3	60	100
5	85.7	200	166.7	140	25
6	114.3	125	133.3	120	100
7	57.1	100	100	100	125
8	142.9	25	33.3	80	200
9	171.4	100	66.7	100	50
10	85.7	125	133.3	120	75
11	114.3	50	33.3	60	125
12	57.1	75	66.7	80	150

ثالثاً: نجد معدل كل فصل لكافة السنوات.

أي في مثالنا نجد معدل مبيعات شهر (1) للسنوات 1997-1993 يكون:

$$\frac{114.3 + 150 + 166.7 + 140 + 75}{5}$$

$$= \frac{646}{5}$$

$$= 129.2$$

والجدول التالي يبين مجموع الأرقام القياسية الشهرية لكل شهر من أشهر السنة للسنوات المختلفة ومعدلاتها:

الشهر	المجموع الشهري	المعدل
1	646	129.20
2	452.1	90.42
3	518.1	103.62
4	371.9	74.38
5	617.4	123.48
6	592.6	118.52
7	482.1	96.42
8	481.2	96.24
9	488.2	97.62
10	539	107.80
11	382.6	76.52
12	428.8	85.76
المجموع		1199.98

ملاحظات:

1- يجب أن يكون مجموع المعدلات الفصلية للسنوات المختلفة مساوياً عدد الفصول × 100 وفي مثالنا هذا كان عدد الفصول في السنة "عدد الأشهر" (12) فصلاً وبالتالي يجب أن يكون مجموع العمود الأخير =1200.

2- إذا اختلف مجموع المعدلات الفصلية للسنوات عن عدد الفصول × 100 يكون السبب في ذلك راجعاً إلى مركبة الخطأ. وإذا كان الاختلاف كبيراً يكون تأثير مركبة الخطأ كبيراً ولذلك يجب تعديل المعدلات الفصلية بضرب كل معدل في المقدار.

$$\frac{\text{عدد الفصول} \times 100}{\text{مجموع المعدلات الفصلية}}$$

وذلك حتى نقلل من تأثير مركبة الخطأ.

ففي مثالنا السابق نلاحظ أن الفرق بين 98. 1199، 1200، هو 0.02 وهو قيمة صغيرة جداً، أي يمكن إهمال تأثير مركبة الخطأ.

تمارين

1- للسلسلة الزمنية التالية:

30، 39، 45 ، 47، 46 ، 43 ، 69 ، 57 ، 61 ، 67

أوجد ما يلي:

a) معامل الخشونة لهذه السلسلة.

b) سلسلة المعدلات المتحركة بطول 4.

c) معامل الخشونة للسلسلة الناتجة في b.

2- الجدول التالي يمثل الكميات المصدرة لسلعة ما مقدرة بآلاف الدنانير، للفترة (1993-2002).

السنة	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
الكمية المصدرة بآلاف الدنانير	135	146	153	163	178	196	200	170	226	195

أوجد معادلة الاتجاه العام. بالطرق التالية:

a) المربعات الصغرى.

b) التمهيد باليد.

c) نصف السلسلة.

d) المتوسطات المتحركة.

3- ما المقصود بما يلي:

1) السلسلة الزمنية.

2) معامل الخشونة.

3) تحليل السلسلة الزمنية.

4) مركبة الاتجاه العام.

4- الجدول التالي يمثل أرباح متجر ما بآلاف الدنانير في الفترة (1998-2005).

السنة	1998	1999	2000	2001	2002	2003	2004	2005
الأرباح	70	75	78	83	85	81	89	80

أوجد ما يلي:

a) معادلة الاتجاه العام بطريقة المربعات الصغرى، ثم قدر الأرباح في الفترة (1998-2005).

b) قدر أرباح هذا المتجر في سنة 2007.

5- في السؤال السابق ارسم المنحنى التاريخي للسلسلة ثم قدر الأرباح بطريقة المتوسطات المتحركة.

6- إذا علمت أن القيمة الحقيقية لظاهرة ما هي (8) في عام 1995، 6 في عام 1998، وإذا كانت القيمة المقدرة في هذه السنوات هي 7.7 ، 5.7 على التوالي، فما هي معادلة الاتجاه العام للفترة (1991-2000)؟

7- سلسلة زمنية عدد عناصرها (157)، جد عدد الأوساط المتحركة بطول (15)؟

8- الجدول التالي يبين معدل درجات الحرارة للفصول الأربعة للسنوات 1991-2000.

السنة / الفصل	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**
Spring	20	22	25	21	24	23	25	20
summer	35	32	40	33	30	37	34	38
Fall	22	20	24	27	23	30	25	26
winter	4	6	10	7	5	3	12	8

حلل مركبة الفصل باستخدام طريقة النسب المئوية للمعدل.

9- إذا كانت معدلات الأرقام القياسية الفصلية لمركبة الفصل للفصول الأربعة للسنوات (1985-1995) معطاة بالجدول التالي:

winter	Fall	summer	spring	الفصل
85	125	90	150	المعدل

قلل من تأثير مركبة الخطأ. على هذه المعدلات؟

10-الجدولان التاليان يبينان المعدل الشهري لإنتاج مصنع بآلاف القطع للسنوات (1989 – 1996)..

السنوات / الشهر	1989	1990	1991	1992
كانون الثاني	55	85	90	120
شباط	60	90	105	110
آذار	70	95	100	150
نيسان	65	120	130	140
أيار	80	100	120	140
حزيران	100	85	70	120
تموز	95	125	135	155
آب	105	120	140	150
أيلول	110	95	120	100
تشرين أول	100	60	70	65
تشرين ثاني	90	100	120	140
كانون أول	85	80	115	130

السنوات / الشهر	1993	1994	1995	1996
كانون الثاني	88	150	125	200
شباط	170	130	98	185
آذار	140	100	85	140
نيسان	135	150	140	200
أيار	190	110	123	170
حزيران	135	160	155	190
تموز	200	165	160	200
آب	195	170	165	195
أيلول	85	125	120	95
تشرين أول	80	100	95	105
تشرين ثاني	90	110	200	190
كانون أول	105	100	150	170

حلل مركبة الفصل باستخدام طريقة النسب المئوية للمعدل.

ملحـــق

(1)

رمز المجموع

(Sigma Notation)

تعريف:

إذا كان f(x) اقتران فيمكن التعبير عن

$$f(1) + f(2) + \ldots + f(n)$$

بالرمز

$$\sum_{r=1}^{n} f(r)$$

مثال:

اكتب ما يلي باستخدام رمز المجموع

a) $2 + 4 + \ldots + 20$

b) $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{128}$

c) $-1 + 1 - 1 + 1 - 1 + 1 - 1 + 1$

الحل:

a) $\sum_{r=1}^{10} 2r$

b) $\sum_{r=1}^{7} \left(\frac{1}{2}\right)^r$

c) $\sum_{r=1}^{8} (-1)^r$

مثال:

جد قيمة $\sum_{r=1}^{4} r^2$

الحل:

$$\sum_{r=1}^{4} r^2 = 1^2 + 2^2 + 3^2 + 4^2$$

$$= 30$$

خصائص رمز المجموع

1) $$\sum_{r=1}^{n}\left[f(r) m\, g(r)\right] = \sum_{r=1}^{n} f(r) m \sum_{r=1}^{n} g(r)$$

2) $$\sum_{r=1}^{n} a\, f(r) = a\sum_{r=1}^{n} f(r)$$

3) $$\sum_{r=m}^{n} f(r) = \sum_{r=1}^{n} f(r) - \sum_{r=1}^{m-1} f(r)$$

4) $$\sum_{r=m}^{n} f(r) = \sum_{r=m-s}^{n-s} f(r+s) - \sum_{r=m+s}^{n-s} f(r-s)$$

5) $$\sum_{r=1}^{n} a = na$$

6) $$\sum_{r=m}^{n} a = (n-m+1)a$$

مبدأ العد

إذا أمكن إجراء عملية على مرحلتين بحيث تتم الأولى بـ (n) من الطرق والثانية بـ (m)من الطرق فإن:

1- العملية تتم بـ (n×m) من الطرق.

2- المرحلة الأولى أو الثانية تتم بـ (n+m) من الطرق.

مثال:

محل تجاري لديه ثلاث أنواع من القمصان وأربعة أنواع من البنطلونات فبكم طريقة يمكن أن يختار شخص:

a) قميصاً وبنطلوناً.

b) قميصا أو بنطلوناً.

الحل:

a) $3\times4 = 12$ طريقة.

b) $3 + 4 = 7$ طرق

مضروب العدد والتوافيق

تعريف: إذا كان ن عدد طبيعياً فإن

a) مضروب العدد n هو $n! = n(n-1)\,(n-2)\,(n-3)\,\ldots\,(2)\,(1)$

b) $0! = 1$

مثال:

a) $5!$ b) $\frac{9!}{7!}$ c) $(5+3)\,!$ d) $5! + 3!$

الحل:

a) $5! = 5 \times 4 \times 3 \times 2 \times 1$

b) $\frac{9!}{7!} = \frac{9 \times 8 \times 7!}{7!}$

$= 72$

c) $(5+3)! = 8! = 8\times7\times6\times5\times4\times3\times2\times1$

$= 40320$

d) $5! + 3! = 120+6$

$= 126$

تعريف: توافيق n من الأشياء مأخوذة r في كل مرة يرمز له بالرمز

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \qquad r \le n$$

مثال: جد

a) $\binom{5}{2}$ b) $\binom{7}{3}$

الحل:

a) $\binom{5}{2} = \frac{5!}{2! \times 3!}$

$=10$

b) $\binom{7}{3} = \frac{7!}{3! \times 4!} = 35$

خصائص التوافيق

1) $\binom{n}{0} = 1 \ , \ \binom{n}{n} = 1$

2) $\binom{n}{1} = n \ , \ \binom{n}{n-1} = n$

3) $\binom{n}{a} = \binom{n}{b} \Leftrightarrow a = b \text{ or } a + b = n$

مثال:

جد حل المعادلة

$$\binom{5}{4} = \binom{5}{x^2}$$

الحل:

$\chi^2 = 4$

$$\Rightarrow x = \pm 2$$

$$\chi^2 + 4 = 5 \quad \Rightarrow \quad x^2 = 1$$

$$\Rightarrow \quad x = \pm 1$$

الوسط التوافقي والوسط الهندسي

تعريف: إذا كانت $x_1, x_2, \ldots, x_n$ مجموعة من المشاهدات فإن:

a) الوسط الهندسي لهذه المشاهدات

$$\overline{X}_g = \sqrt[n]{x_1 x_2 \ldots x_n}$$

b) الوسط التوافقي لهذه المشاهدات

$$\overline{X}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

$$= \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

ملاحظة: باستخدام قوانين اللوغريتمات يمكن كتابة العلاقة الموجودة في الفرع a من التعريف على الصورة

$$Ln\,\overline{X}_g = \frac{1}{n} \sum_{i=1}^{n} \ln x_i$$

مثال:

جد الوسط الهندسي والوسط التوافقي للمشاهدات

8 ، 4 ، 2

الحل:

$$\overline{X}g = \sqrt[3]{2 \times 4 \times 8}$$

$$= 4$$

$$\overline{X}_h = \frac{3}{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}}$$

$$= \overline{X}g = \sqrt[3]{2\times4\times8}$$

$$= 3.43$$

العمليات على المجموعات وقوانينها

إذا كانت A ، B مجموعتين فإن

a) $A \cup B = \{x: x \in A \quad \text{or} \quad x \in B\}$

b) $A \cap B = \{x: x \in A \text{ and } x \in B\}$

c) $A - B = \{x: x \in A \text{ and } x \notin B\}$

d) $\overline{A} = U - A$

حيث U هي المجموعة الكلية وهي أكبر مجموعة قيد الدراسة.

تعريف:

a) A مجموعة جزئية من المجموعة B ويرمز لذلك بالرمز ($A \subset B$)إذا كان كـل عنصر في A موجودا في B.

b) $A = B \Leftrightarrow A \subset B \text{ and } B \subset A$

c) المجموعة الخالية هي تلك المجموعة التي لا يوجد فيها أي عنصر ويرمـز لها بالرمز ϕ أو { }.

مثال:

$U = \{1, 2, 3, 4, 5, 6\}$

$A = \{1,2,4\}$, $B = \{4,5,6\}$

أوجد

1) $A \cup B$ 2) $A \cap B$ 3) $A - B$

4) $\overline{A}$ 5) $\overline{A} \cap \overline{B}$ 6) $\overline{A \cap B}$

الحل:

1) $A \cup B = \{1, 2, 4, 5, 6\}$

2) $A \cap B = \{4\}$

3) $A - B = \{1,2\}$

4) $\overline{A} = \{3,5,6\}$

5) $\overline{A} \cap \overline{B} = \{3,5,6\} \cap \{1,2,3\}$

$= \{3\}$

6) $\overline{A \cap B} = \{\ \}$

أهم قوانين المجموعات

1) $A \cup A = A$
2) $A \cap A = A$
3) $A \cup B = B \cup A$
4) $A \cap B = B \cap A$
5) $A \cap (B \cap C) = (A \cap B) \cap C$
6) $A \cup (B \cup C) = (A \cup B) \cup C$

نظرية:

إذا كانت A ، Bمجموعتين ، $A \subset B$ فإن

a) $A \cap B = A$

b) $A \cup B = B$

تعريف: مجموعة القوى للمجموعةA يرمز لها بالرمز P(A) وتكون

$P(A) = \{X : X \subset A\}$

ويكون

$\#(P(A)) = 2^{\#(A)}$

مثال:

إذا كانت $A= \{1,2,3\}$

الحل:

$$P(A) = \{\phi , \{1\} ,\{2\} , \{3\} , \{1,2\} , \{1,3\}, \{2,3\} , A\}$$

مجموعات الأعداد

1- مجموعة الأعداد الطبيعية $\mathbf{N}= \{1,2,3,\ldots\}$

2- مجموعة الأعداد الصحيحة $\mathbf{Z}= \{\ldots,-2,-1,0,1,2,\ldots\}$

3- مجموعة الأعداد الحقيقية $\mathbf{R}= (-\infty,\infty)$

ملحـــق

(2)

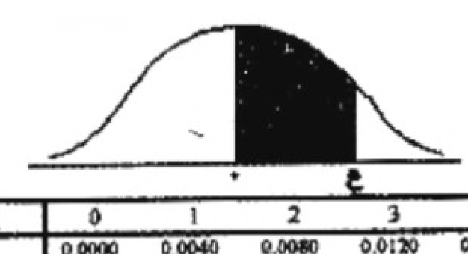
جدول الأرقام العشوائية

51772	74640	42331	29044	46621	62898	93582	04186	19640	87056
54033	23491	83587	06568	21960	21387	76105	10863	97453	90581
45939	60173	52078	25424	11645	55870	56974	37428	93507	94271
03585	79353	81938	82322	96799	85659	36081	50884	14070	42187
64937	03355	95863	20790	65304	55189	00745	62528	11822	15804
15630	64759	51135	98527	62586	41889	25439	88036	24034	67283
21681	91157	77331	60710	52290	16835	48653	71590	16159	14676
91097	17480	29414	06829	87843	28195	27279	47152	35683	47280
50532	25496	95652	42457	78547	76552	50020	24849	52984	76168
07136	40876	79971	54195	25708	51817	36732	72484	94923	75985
27989	64728	10744	089396	56242	90985	28868	99431	50995	20507
86181	78949	86601	46258	00477	25234	09903	36574	72139	70185
54308	21154	97810	86764	82869	11785	55261	59009	38714	38723
65541	34371	09591	07889	58892	92843	72828	91341	84821	63886
08263	65952	85762	64236	39238	18776	84303	99247	46149	03229
39817	67906	48236	16057	81812	15815	63700	85915	19219	45943
62257	04077	79443	95203	02479	30763	92486	54083	23631	05325
53298	90276	62545	21944	16580	03878	07516	95715	02526	33537

أخذت الجداول في هذا الملحق من كتاب

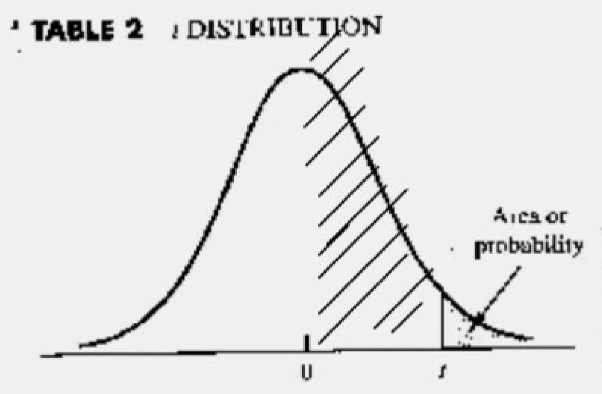
(Theory and Problems of Statistics) By (Murray R. Spiegel)

جدول التوزيع الطبيعي المعياري


ع

ع	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4973	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

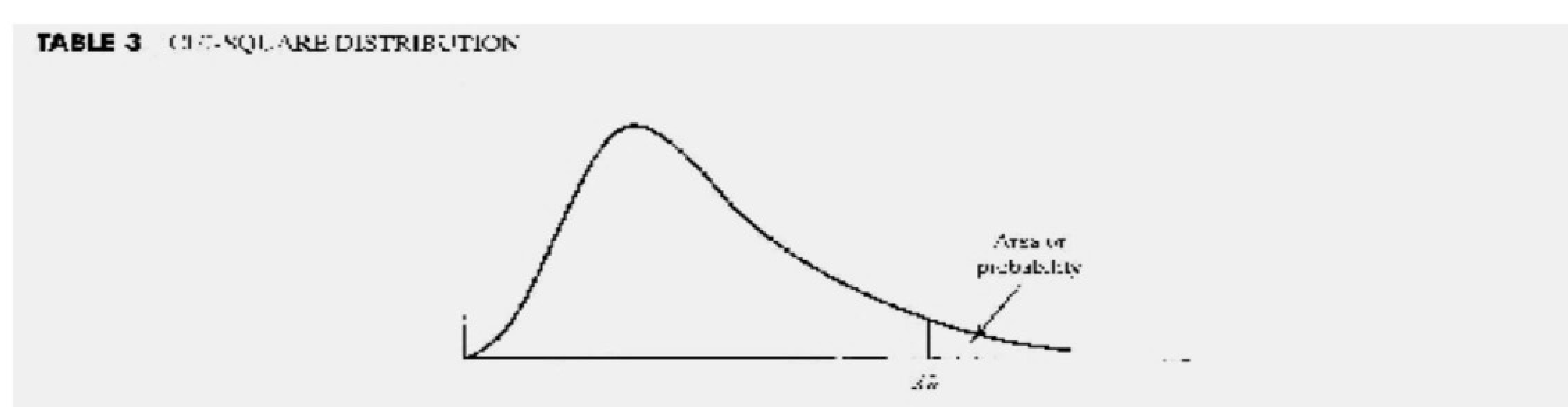
TABLE 2 t DISTRIBUTION



Entries in the table give t values for an area or probability in the upper tail of the t distribution. For example, with 10 degrees of freedom and a .05 area in the upper tail, $t_{.05} = 1.812$.

Degrees of Freedom	Area in Upper Tail .10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

TABLE 3 CHI-SQUARE DISTRIBUTION



Entries in the table give χ^2_α values, where α is the area or probability in the upper tail of the chi-square distribution. For example, with 10 degrees of freedom and .01 area in the upper tail, $\chi^2_{.01} = 23.2093$.

Degrees of Freedom	Area in Upper Tail									
	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	$392,704 \times 10^{-10}$	$157,088 \times 10^{-9}$	$982,069 \times 10^{-9}$	$393,214 \times 10^{-8}$	.0157908	2.70554	3.84146	5.02389	6.63490	7.87944
2	.0100251	.0201007	.0506356	.102587	.210720	4.60517	5.99147	7.37776	9.21034	10.5966
3	.0717212	.114832	.215795	.351846	.584375	6.25139	7.81473	9.34840	11.3449	12.8381
4	.206990	.297110	.484419	.710721	1.063623	7.77944	9.48773	11.1433	13.2767	14.8602
5	.411740	.554300	.831211	1.145476	1.61031	9.23635	11.0705	12.8325	15.0863	16.7496
6	.675727	.872085	1.237347	1.63539	2.20413	10.6446	12.5916	14.4494	16.8119	18.5476
7	.989265	1.239043	1.68987	2.16735	2.83311	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.344419	1.646482	2.17973	2.73264	3.48954	13.3616	15.5073	17.5346	20.0902	21.9550
9	1.734926	2.087912	2.70039	3.32511	4.16816	14.6837	16.9190	19.0228	21.6660	23.5893
10	2.15585	2.55821	3.24697	3.94030	4.86518	15.9871	18.3070	20.4831	23.2093	25.1882
11	2.60321	3.05347	3.81575	4.57481	5.57779	17.2750	19.6751	21.9200	24.7250	26.7569
12	3.07382	3.57056	4.40379	5.22603	6.30380	18.5494	21.0261	23.3367	26.2170	28.2995
13	3.56503	4.10691	5.00874	5.89186	7.04150	19.8119	22.3621	24.7356	27.6883	29.8194
14	4.07468	4.66043	5.62872	6.57063	7.78953	21.0642	23.6848	26.1190	29.1413	31.3193
15	4.60094	5.22935	6.26214	7.26094	8.54675	22.3072	24.9958	27.4884	30.5779	32.8013
16	5.14224	5.81221	6.90766	7.96164	9.31223	23.5418	26.2962	28.8454	31.9999	34.2672
17	5.69724	6.40776	7.56418	8.67176	10.0852	24.7690	27.5871	30.1910	33.4087	35.7185
18	6.26481	7.01491	8.23075	9.39046	10.8649	25.9894	28.8693	31.5264	34.8053	37.1564
19	6.84398	7.63273	8.90655	10.1170	11.6509	27.2036	30.1435	32.8523	36.1908	38.5822

20	7.43386	8.26040	9.59083	10.8508	12.4426	28.4120	31.4104	34.1696	37.5662	39.9968
21	8.03366	8.89720	10.28293	11.5913	13.2396	29.6151	32.6705	35.4789	38.9321	41.4010
22	8.64272	9.54249	10.9823	12.3380	14.0415	30.8133	33.9244	36.7807	40.2894	42.7958
23	9.26042	10.19567	11.6885	13.0905	14.8479	32.0069	35.1725	38.0757	41.6384	44.1813
24	9.88623	10.8564	12.4011	13.8484	15.6587	33.1963	36.4151	39.3641	42.9798	45.5585
25	10.5197	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3141	46.9278
26	11.1603	12.1981	13.8439	15.3791	17.2919	35.5631	38.8852	41.9232	45.6417	48.2899
27	11.8076	12.8786	14.5733	16.1513	18.1138	36.7412	40.1133	43.1944	46.9630	49.6449
28	12.4613	13.5648	15.3079	16.9279	18.9392	37.9159	41.3372	44.4607	48.2782	50.9933
29	13.1211	14.2565	16.0471	17.7083	19.7677	39.0875	42.5569	45.7222	49.5879	52.3356
30	13.7867	14.9535	16.7908	18.4926	20.5992	40.2560	43.7729	46.9792	50.8922	53.6720
40	20.7065	22.1643	24.4331	26.5093	29.0505	51.8050	55.7585	59.3417	63.6907	66.7659
50	27.9907	29.7067	32.3574	34.7642	37.6886	63.1671	67.5048	71.4202	76.1539	79.4900
60	35.5346	37.4848	40.4817	43.1879	46.4589	74.3970	79.0819	83.2976	88.3794	91.9517
70	43.2752	45.4418	48.7576	51.7393	55.3290	85.5271	90.5312	95.0231	100.425	104.215
80	51.1720	53.5400	57.1532	60.3915	64.2778	96.5782	101.879	106.629	112.329	116.321
90	59.1963	61.7541	65.6466	69.1260	73.2912	107.565	113.145	118.136	124.116	128.299
100	67.3276	70.0648	74.2219	77.9295	82.3581	118.498	124.342	129.561	135.807	140.169

This table is reprinted by permission of Oxford University Press on behalf of The Biometrika Trustees from Table 8, Percentage Points of the χ^2 Distribution, by E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., 1966.

فهرسة المصطلحات عربي إنجليزي

Index

Natural numbers	الأعداد الطبيعية
Probality density function	اقتران الكثافة الاحتمالية
Mean deviation	الانحراف المتوسط
Standard deviation	الانحراف المعياري
Variance	التباين
Random experiment	تجربة عشوائية
Kurtosis	التفرطح
Intersection	التقاطع
Estimation of trend	تقدير مركبة الاتجاه
Estimation of seasonal variation	تقدير مركبة الفصل
Frequency	التكرار
Cumulative frequency	- التكرار التراكمي
Relative frequency	- التكرار النسبي
Frequency histogram	- المدرج التكراري
Frequency polygon	- المضلع التكراري
Frequency curve	- المنحنى التكراري
Cumulative frequency curve	- المنحنى التكراري التراكمي
Relative frequency curve	- المنحنى التكراري النسبي
Forecast	التنبؤ
Expectation	التوقع
Combinations	التوافيق
Frequency distribution	التوزيع التكراري
Frequency distribution table	جدول التوزيع التكراري
Probablity distribution	التوزيع الاحتمالي

Mono-mode distribution	توزيع أحادي المنوال
Mulry-mode distribution	- توزيع متعدد المنوالات
Normal distribution	توزيع طبيعي
Standared normal distribution	- توزيع طبيعي معياري
Skewed distribution	توزيع ملتو
Event	الحادث
Sure event	- الحادث الأكيد
Null event	- الحادث المستحيل
Independent events	- الحوادث المستقلة
Simple event	- الحادث البسيط
Complementary event	- الحادث المتمم
Compound events	- الحادث المركب
Disjoint events	- الحوادث المنفصلة
Mutually exclusive events	- الحوادث المتباعدة والشاملة
Lower class limit	الحد الأدنى للفئة
Upper class event	الحد الأعلى للفئة
Lower class boundary	الحد الأدنى الفعلي للفئة
Upper class boundary	الحد الأعلى الفعلي للفئة
Sample size	حجم العينة
Population size	حجم المجتمع
Actual income	الدخل الحقيقي
Quartile	الربيع
Range of the variable	رتبة المتغير
Index number	الرقم القياسي

- الرقم القياسي البسيط	Simple index number
- الرقم القياسي التجميعي البسيط	Simple aggregate index number
- الرقم القياسي النسبي البسيط	Simple relative index number
- الرقم القياسي للأسعار	Price index number
- الرقم القياسي للكميات	Quantities index number
- الرقم القياسي لتكاليف المعيشة	Cost of livint index number
- الرقم القياسي المرجح	Weighted index number
- الرقم القياسي العضلي	Seasonal index number
- رقم باش للأسعار	Paasche index number for prices
- رقم باش للكميات	Paasche index number for quantities
- رقم لاسبير للأسعار	Laspeyre index number for prices
- رقم لاسبير للكميات	Laspeyre index number for quantities
- رقم فشر الأمثل للأسعار	Fisher's ideal index number for prices
- رقم فشر الأمثل للكميات	Fisher's ideal index number for quantities
- رقم مارشال للأسعار	Marshall index number for prices
- رقم مارشال للكميات	Marshall index number for quantities
رمز المجموع	Sum simple
زمن الأساس	Base time
زمن المقارنة	Comparison time
السلسلة الزمنية	Time series
سن الحل	Pregnancy rage
طريقة التمهيد باليد	Free hand method
طريقة الخط المنكسر	Broken line method

Smooth line method	الخط المنحني
Picture method	طريقة الصور
Pie charts method	طريقة القطاع الدائري
Moving averages method	طريقة المتوسطات المتحركة
Rectangular method	طريقة المستطيلات
The method of least squers	طريقة المربعات الصغرى
The average percentage method	طريقة النسب المئوية للمعدل
Ratio to trend method	طريقة النسبة إلى الاتجاه العام
Ratio to moving average method	طريقة النسبة إلى المتوسطات المتحركة
The method of semi-averages	طريقة نصف السلسلة
The link relative method	طريقة الوصل النسبية
Moments	العزوم
Deciles	العشير
Positive relation	علاقة إيجابية (طردية)
Negative relation	علاقة سلبية (عكسية)
Raw mark	العلاقة الخام
Standared mark	العلاقة المعيارية
Sample	العينة
Random sample	- العينة العشوائية
Stratified sample	- العينة الطبقية
Systematic random	- العينة العشوائية المنتظمة
Standared sample	- العينة المعيارية
Class	الفئة
Class strength	- طول الفئة

– مركز الفئة	Class mark
– الفئة المئينية	Percentril class
– الفئة المنوالية	Mode class
– الفئة الوسيطية	Mediar class
فضاء العينة	Sample space
قوانين ديمورغان	Demorgan's laws
الالتواء	Skewness
المئين	Percentile
– الرتبة المئينية	Percentile rank
مبدأ العد	Counting principle
المتغير العشوائي	Random variable
– مدى المتغير العشوائي	Rane of random variable
– المتغير العشوائي المتصل	Continuous random variable
– المتغير العشوائي المنفصل	Discrete random variable
متممة	Complement
المتوسطات المتحركة	Moving averages
مجموعة	Set
– مجموعة خالية	Empty set
– مجموعة جزئية	Subset
– مجموعة كلية	Universal set
العمليات على المجموعات	Operations of sets
المدى (المدى المطلق)	Range
المدى الربيعي	Interquartile range
المدى العشيري	Deciles range

Component of time series	مركبة السلسلة الزمنية
Secular trend	مركبة الاتجاه
Cyclical component	مركبة الدورة
Irregular (random) variation	مركبة الخطأ (العشوائية)
Seasonal variation	مركبة الفصل
Observation	المشاهدة (المفردة)
Factorial	المضروب
Equation of an appropriate trend	معادلة الاتجاه العام
Regression line equation	معادلة خط الانحدار
Roughness coefficient	معامل الخشونة
Crude partiality rate	معدل الخصوبة العام
Married woman's partiality rate	معدل الخصوبة للنساء المتزوجات
Age specific vertility rate	معدل الخصوبة المحدد بفئة عمرية
Seasonal average	المعدل الفصلي
Immigration rate	معدل الهجرة
Crude death rate	معدل الوفاة الخام
Infant nortality rate	معدل وفيات الرضع
Neonatal mortality rate	معدل وفيات حديثي الولادة
Maternal nortality rate	معدل وفيات الأمومة
Crude birth rate	معدل الولادة الخام
Measures of dispersion	مقاييس التشتت
Central tendency measures	مقاييس النزعة المركزية
Historical curve	المنحنى التاريخي
Mode	المنوال

New borns	المواليد الأحياء
Binomial theorem	نظرية ذات الحدين
Bayes theorem	نظرية بيز
Immigration	الهجرة
Mean	الوسط الحسابي
Harmonic mean	- الوسط التوافقي
Guessed mean	- الوسط الفرضي
Weighted mean	-الوسط الموزون
Geometric	- الوسط الهندسي
Median	الوسيط

المراجع

أولا: المراجع العربية

1- مختار الهانسي، مبادئ الإحصاء، الدار الجامعية، بيروت 1993.

2- محمد أبو صالح، مقدمة في الإحصاء، مركز الكتب الأردني، عمان 1990.

3- محمد مظلوم حمدي، طرق الإحصاء، دار المعارف مصر، 1965.

4- أحمد عبادة سرحان، الإحصاء، مؤسسة شباب الجامعة، الإسكندرية 1977.

5- عبد الحسين زيني، الإحصاء السكاني، وزارة التعليم العالي، بغداد 1980.

6- مدني دسوقي مصطفى، مبادئ في علم الإحصاء، دار النهضة العربية، مصر 1977.

ثانيا: المراجع الأجنبية

1- Murray R. Spiegel, Theory and problems of Statistics, Mc Graw-Hill. New York 1987.

2- William Mendenhall, Introduction to Probability and Statistics, 5 the edition.

3- David R. Anderson, Statistics for business and economics, 8th edition, south-western, a division of Thomas learning.